GENERAL MANAGEMENT ASSISTANCE AGREEMENT (GMAC)
Agreement No: 674-C-00-01-10051-00

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

This report was prepared under Mega-Tech, Inc.'s prime agreement with USAID and addresses USAID/South Africa's Strategic Objective No. 6: Increased access to shelter and environmentally sound municipal services

Please direct all queries regarding this report to:

Khulisa Management Services
26 Seventh Avenue
Parktown North 2196
SOUTH AFRICA
Tel: +27 (11) 447-6464
Fax: +27 (11) 447-6468
Email: info@khulisa.com

OR

MegaTech/South Africa
General Management and Assistance Agreement Office
Bank Forum Building
Lobby 1, Second Floor
337 Veale Street
New Muckleneuk
0181 Pretoria RSA
Tel. (012) 452-0060
Fax (012) 452-0070
Email megatech@intekom.co.za

Or

Mega-Tech, Inc.
180 South Washington Street, Suite 200
Falls Church, VA 22046
Tel. (703) 534-1629
Fax (703) 534-7208
Email info@mgtech-world.com

Table of Contents

Figures

Appendices

Acronyms

AE	Agama Energy
CMDA	Cato Manor Development Association
COP	Chief of Party
CT	Cape Technikon
DEAT	Department of Environmental Affairs and Tourism
DQA	Data Quality Assessment
FTFA	Food and Trees for Africa
FRB	First Rand Bank
GAAP	Generally Accepted Accounting Principles
GCC	Global Climate Change
HIV/AIDS	Human Immunodeficiency Virus / Auto Immune Deficiency Syndrome
IIEC	International Institute for Energy Conservation
IPD	Isandla Partners in Development
ISO	International organization for Standardization
JHC	Johannesburg Housing Company
JMC	Johannesburg Metropolitan Council
KCIHT	Kutlwanong Civic Integrated Housing Trust
LDF	Lynedoch Development Foundation
ME	Midrand Ecocity
MIIU	Municipal Infrastructure Investment Unit
NDISWE	National Development Initiative for Social Welfare
NLP	Ndlandlamuka Local Project
OF	Organizational Framework
PHPT	Peoples Housing Partnership Trust
PMP	Performance Monitoring Plan
PPTKN	Kwa-Zulu Natal Project Preparation Trust
PWC	Price Waterhouse & Coopers
RWD	Chemonics
SDF	Soweto Development Foundation
SO6	Strategic Objective Six
SOW	Scope of work
TIPS	Performance Monitoring and Evaluation Guidelines for Indicators and Data Quality
UCT	University of Cape Town
USAID	United States Agency for International Development

1. Executive Summary

The primary objective of this DQA was to assess the quality of the SO6 indicators and the quality of the data collected and reported on by individual implementing partners. It is expected that this assessment will support the SO6 team in its efforts to strengthen their Performance Monitoring Plan (PMP).

The methodological approach was dictated to some degree by the list of partners supplied to the Team. The decision was taken, and agreed to by USAID/SA, that the DQA would be based on Probability-Based Sampling of the given population, understanding that inherent limitations exist with sampling a pre-existing sample. The given list of partners was subdivided into three Populations – those that report data directly to USAID/SA, those that report data through a contractor and those for which data would be reviewed at USAID/SA.

In brief the activities performed for this DQA were as follows:

a. Preliminary review of the DQA documentation (ADS guidelines, TIPS, PWC Toolkit, etc), and preliminary discussion with Mega-Tech to identify priority issues & concerns.

b. Initial consultations with mission personnel to:
   - Review the overall scope of work for the SO6 DQA,
   - Clarify the set of indicators to be covered; and
   - Discuss mission and bureau issues/concerns about indicators and data quality.

c. Preparation of performance indicator and data quality assessment tools to:
   - Guide the team about indicator and data quality issues, and
   - Provide the information source for indicator and data quality assessment tables.

d. Intensive consultations with implementing partners, where possible, to:
- Conduct a validation exercise of data quality with each partner based on the information contained in Data Quality Assessment Checklist using the ISO 19011 audit guidelines method, and
- Identify data strengths, vulnerabilities and risks.

e. Quality assurance and quality control reviews of incoming data validations.

The Team's approach to assessing the quality of data collected and reported on by the individual partners was based on the internationally recognized "Guidelines for Quality and/or Environmental Management Systems Auditing". The results of the validation process allowed the Team to assess each organization's capacity to collect and report on SO6 indicator data, and to point out strengths and vulnerabilities in the partners' data systems.

Audit findings are an evaluation of the collected audit evidence against the audit criteria and were defined as follows:

- A **NON-CONFORMITY** was declared when the audit evidence showed that there had been non-fulfillment of a criterion. Such non-conformities were classified as **MINOR** (despite a failure, the overall data quality characteristic e.g. validity, could still be achieved) or **MAJOR** (a failure which prevented the achievement of the overall data quality characteristic).

- An **OBSERVATION** was noted when a **STRENGTH** or **VULNERABILITY** was noted in the partner's data quality system, which importantly, was not a non-conformity.
- **Strengths** are identified with the purpose of giving positive feedback.
- **Vulnerabilities** are identified with the express purpose of giving an indication of risk and extremely useful for internal management use.

Achievement of the criteria by the partners was scored in order to give an overall assessment of the partners' conformance to the data quality requirements. The scoring rubric for each item in the data quality check list was as follows:

- If the partner met the criterion in its stated form it received a score of three
- If a minor non-conformity had been identified it received a score of two
- If a major non-conformity had been identified, it received a score of one.

The results of the DQA, per quality attribute, were averaged over the number of required criteria. The results are presented as a nominal scale between one and three where one indicates an absolute failure to achieve the required attribute and three indicates that ideal data quality is achieved.

By assessing the data quality associated with each indicator, per partner, this DQA has evaluated the ability of each of the partners to meet the USAID basic data quality requirements. Summation of the results for the partner organizations, per performance indicator, are given below.

1.1. Summation of DQA Results

Population One

A total of nine partners were reviewed in population one. Of these partners, one audit was declared not feasible due to the absence of any auditable information on the day of audit.

The average scores per quality characteristic are presented in Figure E1. This clearly demonstrates that the risk presented to USAID in terms of data quality is least in terms of the integrity of the data, i.e. it can be demonstrated that the data is free of undue or unethical manipulation. Greatest risk is presented in terms of the precision of data where in many cases there are high levels of error inherent in the data. This area also shows the greatest variability amongst partners with scores ranging from the minimum of 1 to the maximum of 3. The low scores attained by some partners with regards to data validity occur when data is 'tortured' in an effort to make it match the indicator reported on. This is a practice which must be actively discouraged by USAID/SA.

Where reliability is low this is usually due to variations in collection methodologies and the difficulty of achieving repeatable results. Low scores for timeliness are most often the result of the reporting of data being inadequate in terms of allowing USAID to identify or manage any potential data-related risk.

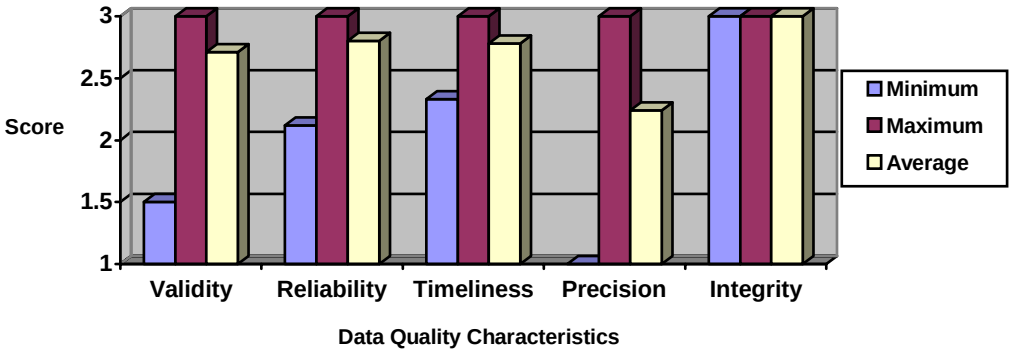

Figure E1: Average scores per quality characteristic for population one.

Population Two

Population two consisted of eight partners who report their data via a GMAC held by Mega-Tech (MT) to USAID. In all cases the partners concerned were reporting GCC data and as the main purpose of the DQA was to assess SO6 data it was considered that on-site visits for these partners was not essential. The assessment is therefore more one of the data quality management systems of Mega-Tech than the data themselves.

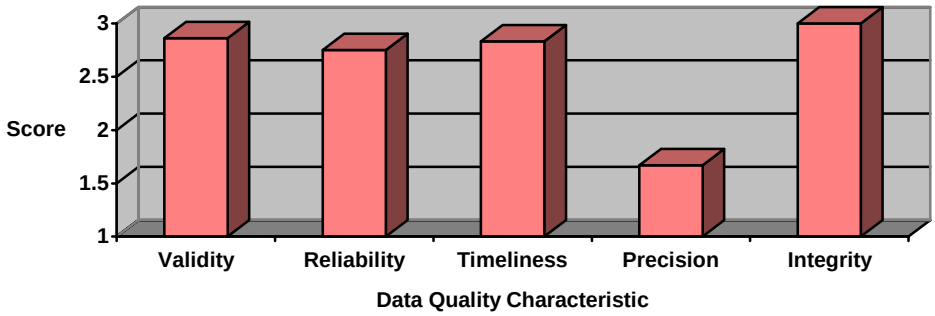

Figure E2: Average scores per quality characteristic for population two.

Risk to USAID/SA

The risks presented by MT do not have their origins at MT – they are systems issues such as the management of indicators, especially the definitions of the GCC indicators. USAID/SA have also made many changes over the past few years. In terms of Mega-Tech data quality systems they demonstrate an overall low risk, with the greatest risk presenting in the area of precision. The results suggest that the handling of partners through a GMAC reduces data quality risk.

Deliverables

MT are responsible for collecting the data required from these contracted partners as well as the data quality of these partners. MT designed a standard report form for the partners contracted through MT – this form was introduced in October/November 2002. Partners attended a workshop in October 2002 where the form was explained to them. MT collates all the reports into one report for USAID/SA. Until recently, the raw data as supplied by partners was forwarded to USAID/SA, “as is”.

On the latest agreements with partners, grantees were given a choice to decide which indicators to report against as per USAID/SA instructions. MT subsequently identified that the indicators reported on were not reflective of the grant programs and brought this to the attention of SO6. After a meeting to discuss the issue, SO6 asked MT to go back to the most recent SO6 (non-GCC) grantees and

renegotiate more relevant indicators, which has been done. MT also designed a checklist – "Data Quality Questionnaire". This questionnaire is to be attached to the data form for the October 2003 data collection for the annual report to Washington. This process has just started so there are no records as yet on this new improved system.

Population Three

Two partners were to be assessed as part of population three. Of these no records could be located for one of the partners. The remaining partner is a borrower under the Housing Guaranty program and had negotiated freedom not to be audited on its data reporting. The DQA was therefore undertaken on the basis of documentary evidence in the SO6 office.

In terms of data quality the results of this DQA are shown below in Figure 3. The low score given to precision is based on the audit being unable to demonstrate the margin of error; likewise this applies for the integrity of the data.

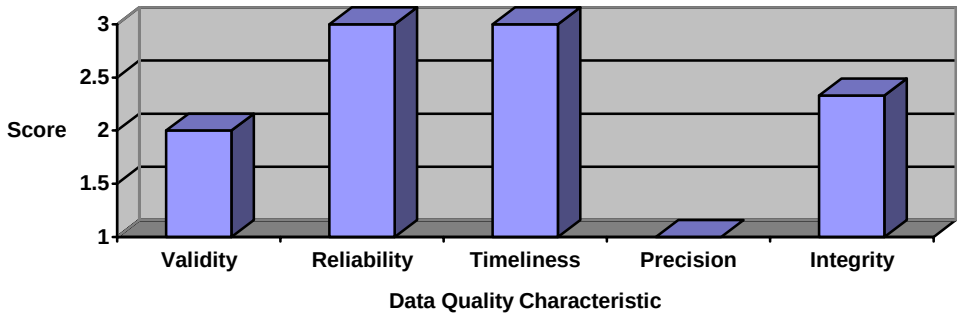

Figure E3: Average scores per quality characteristic for population three.

Summary

1. Systems

There were systems difficulties in two senses:

1.1 Data concerning the cooperating partners appeared to be difficult to access and sometimes was out of date.

1.2 There was not a clear audit trail regarding how the data from the partners was aggregated into the numbers used in the PMP.

2. Lack of clarity regarding the Indicators

Most partners were either not aware of which indicators they were reporting against, or else selected the indicators that they wishes to report against. This reduces the effectiveness of the PMP as a management tool and potentially affects the reliability of the data which may have to be manipulated in order to fit into the specific requirements of the indicator.

3. Precision

There are serious difficulties in respect of precision. These derive mainly from a lack of clarity in terms of the definitions and the need to convert data through formulae that cannot be objectively verified.

1.2. Performance Indicator Quality Assessment

The purpose of this exercise was to review the performance indicators against quality characteristics - directness, objectivity, practicality, and adequacy. In each case, the results of the qualitative assessment are given together with a discussion of data quality issues. In addition, recommendations related to each specific indicator are made, where relevant, and which cover such areas as: proposed

indicator definition, rationale, frequency of data collection, methodology, responsibility for data collection, target, and data limitations.

In brief the findings were:

2.1 There are significant overlaps between the indicators.

2.2 The primary data used has, in most cases, to be manipulated to generate the results used. These data manipulations introduce significant risks in terms of the reliability of the data as they are based on formulae, not actual results.

2.3 The reliability of the numbers generated is compromised by a lack of precision in expressions such as "leveraged through USAID supported programs".

2.4 The indicators which relate to policy change should be in narrative form, possibly supported by data points where appropriate. Attempts to construct indices of, for example, policy reform inevitably rely on subjective criteria – even if constructed by highly experienced consultants. They therefore fail to meet the requirement for an indicator to be able to objectively verifiable.

2.5 It is recommended to reduce the number of indicators and use only those which are precise and objectively verifiable.

1.3. Organizational Framework (OF) Description

In undertaking the assessment of both indicator quality and data quality, the Team considered the development model embodied in the strategic framework, and the relationships between the various levels of interactions, such as impacts, outcomes, outputs, activities and inputs. Knowledge of the model and these relationships helped to clarify the characteristics of the indicators, the nature of the data being collected by partners, and partner reporting responsibilities. An Organizational Framework (OF) is used to present this information. The OF is simply another way of representing the SO6 results framework, albeit one that is "stretched" to include the full results ("chain levels") embodied in the SO (inputs, activities, outputs, outcomes and impacts). The respective chain levels correspond to the results framework, as follows:

- Impact level corresponds to the SO
- Outcome level corresponds to IRs
- Output level corresponds to partner performance indicators
- Activity level corresponds to operational activities of implementing partners
- Inputs correspond to important partner data collection items or "data collection points."

In addition, the OF is divided into two broad categories: the external environment (influenced by factors outside partner activities) and internal environment (the local environment in which program partners operate). Partner data collection and reporting responsibilities are represented by blue ovals.

1.4. Current Indicator Relationships - OF

The current construction of the SO PMP has allowed for the majority of the measurable indicators to lie at the impact level of the SO (10 of 17 indicators). As all 10 of these indicators are constructed as output indicators, which result from the activity of a partner, this relationship is inappropriate. By and large the Rand value based indicators can be objectively verified, however the disaggregation into the various municipal services, as required by Washington, is inappropriate for the South African environment. 'Household' based data relative to services is inevitability speculative as opposed to 'shelter' data, which can be measured quantitatively.

At the Outcome level there are four intermediate results of which one measures impact (IR6.1). This relationship is also inappropriate as the nature of data collection, collation and analysis for the

indictors for this result do not support its being placed at the outcome level. As there are manipulated calculations based on the results and outcomes of other activities this result and its indicators must be a measure at the impact level, which remains separate to partner activities.

The indicators for two of the Intermediate Results (IR6.2 &IR6.3) are an operational repeat of indicators at the SO level due to the discrepancies that exist in the definitional issues. This means that in practice they are interpreted and measured as the same entity. This adds no value to the ultimate measurement of the achievement of the SO.

The last Intermediate Result (IR6.4) is the only one, which is aimed at measuring the environmental component of the SO. It construction and use of terminology makes the indicator associated with the result non-measurable. Furthermore the description given to the indicator is unrelated to the IR.

The Figure E4, which is placed at the end of the Summary, illustrates the proposals made by the consultants, as modified following the verbal presentation of the DQA to the Mission.

Discussion held at the meeting held on 12th September
It was decided that the discussion held at the above meeting, which was attended by the Director, Deputy Director, representatives of the Program Office and most of the SO6 team, should be recorded here.

These decisions were, in brief:
1. The proposed reduction in the number of indicators was welcomed.

2. It was agreed that definitional problems created uncertainty, and that definitions would have to be clarified.

3. The question of integrating several indicators of policy change into a single numerical indicator through a matrix or similar device was discussed at length. It was recognized that the indicators which relied on an aggregation of different data were inherently meaningless and did not withstand close scrutiny. On the other hand it was maintained that there was a limited number of well-defined policy changes – about five – which could be measured. Such policy changes, such as the enactment of specific legislation such as the Community Reinvestment Act, or the Municipal Finance Management Act, are clearly defined and recording such events in numerical form was perfectly valid.

It was agreed that policy change (if defined as described above) could be reflected at the Output level in the chart above, i.e. as one of the fundamental indicators, located underneath the oval "Environment based projects".

It was also agreed that policy change was both an output of the SO, and a factor at the Impact level of the SO.

4. Attempts should be made to refine the data used, for example, in the training indicators, so that they measured change in the performance of either the individuals or the organization for which they worked. It was agreed that this might be undertaken in selected cases on a biannual basis.

5. As noted above the team found that too many of the indicators have been set at the Strategic Objective (SO) level, rather than the intermediate result (IR) level. This weakens the logical relationship between them, and reduces their utility as a management tool. It was agreed that this would be reviewed.

Second Meeting

A second meeting was held on Thursday 23 October.

At this the proposed amendments to the PMP were discussed, together with certain further matters of clarification regarding the report and the draft executive summary.

Issues discussed were are follows:
6. Aggregation of data
It was noted that the problem of disaggregation of data as currently adopted is that it may force manipulation of data. If the data cannot be supplied in its disaggregated form it should not be manipulated to do so. Partners who cannot supply disaggregated data should therefore not be required to do so. Furthermore duplication can occur when data disaggregation takes place.

Similarly, cumulative data will be avoided in future.

7. Prospective data
It was noted that MIIU generates a substantial proportion of the numbers reported. It was therefore important to include their work in the PMP.

It was agreed that if there is a clear distinction between *Actual* and *Forecasted* the risk of misrepresentation is avoided. Both types of data may be reported against the same indicator provided that the distinction is maintained.

8. Use of the terms household and shelter units

It was noted that the objections to the use of the word household concern the lack of a clear definition of the term, and a tendency to use it as a synonym for a house. It was agreed that its use can be continued provided proper definitions[1] were used and could be reported against with accuracy. Otherwise, it was noted that the term shelter units was acceptable, and could be reported on with accuracy as at present – namely that a shelter unit is counted an irrevocable contract for its construction is concluded (i.e. this is the point at which a subsidy is granted). This avoids the verification problem of waiting until a unit is occupied.

9. Modification of IR 6.3 regarding "increased environmentally sound municipal services"

A number of changes were proposed.

9.1 The Title of the IR should be changed to "Increased sustainability of municipal services".

9.2 Reporting should be in narrative form, as the type of work – largely technical assistance and capacity building with municipalities and NGOs, and small pilot projects – did not generate valid numerical data. This does not preclude data being included in narrative reports, but its use is mainly illustrative, rather than auditable hard facts. Thus reports on training events might refer to the number of trainees who participated, but no attempt should be made to aggregate that data with other types of training event.

9.3 Two indicators should be used:

9.3.1 Capacity building initiatives undertaken

9.3.2 Improved policy environment

It was agreed that the latter should refer to milestones in the policy reform process, such as the passage of crucial legislation such as the Community Reinvestment Act, or the Municipal Finance Management Act; but that on balance it was more useful to report on these in narrative terms than numerical ones (as had been proposed at the previous meeting, see paragraph 3 above).

[1] Auditable definitions must be kept as simple as possible and avoid use of adjectives and any other qualifiers which are open to interpretation

2. Background to SO6 Data Quality Assessment

The U.S. Agency for International Development (USAID) requires that all program performance data presented in USAID Mission Annual Reports is valid, complete, accurate and consistent with management needs. In support of this requirement, USAID policy (ADS 203) requires that a Data Quality Assessment (DQA) be performed when establishing indicators that are to be reported on in Annual Reports. Data quality must be reassessed as needed, but no less than once every three years. In accordance with Solicitation No. 0112-0603-SOL-ME8, performance of a DQA for USAID/SA's Strategic Objective No. 6 (SO6), "Increased access to shelter and environmentally sound municipal services", was the major purpose of this exercise. This DQA was defined broadly to also include an assessment of the quality of SO6 and GCC data reported as well as SO6 performance indicators.

Therefore, the primary objective of this DQA was to assess both the quality of SO6 indicators and the quality of the data collected and reported on by individual implementing partners. It is expected that this assessment will support the SO6 team in its efforts to strengthen their Performance Monitoring Plan (PMP). In response to the above referenced solicitation, a three person Team was mobilized over a period of approximately one month to carry out the work. In developing the methodology (described in the following section) for this DQA, the Team was guided principally by:

- ADS 203 (Assessment and Learning);
- Price Waterhouse Coopers (PWC) "Performance Management Toolkit";
- Performance Monitoring and Evaluation (TIPS) Guidelines for Indicator and Data Quality; and
- ISO/DIS 19011 Guidelines for Quality and/or Environmental Management Systems Auditing.

On August 4, 2003, USAID/SA briefed the Team on the rationale and context for undertaking the assessment and clarified and/or confirmed the parameters of the DQA. As a result, the following parameters for the exercise were confirmed as follows:

- The primary purpose for conducting the DQA is in order to allow the SO6 Team to assess for any liabilities, contingent or actual, that may arise due to data quality issues and thus enable them to implement improvements for the management of risk.
- The DQA must also provide the SO6 Team with a systems analysis so that they will be able to correct and/or improve their own data handling activities.
- The DQA must take into account the time-sensitive nature of the exercise, due to a key staff member being about to depart from USAID/SA.
- The assessment will be limited to an approved sample of the partners.
- The assessment includes data generated for the GCC indicators.

3. Methodological Approach

1.5. Introduction

Due to the nature of the SOW the Team believed that there was a need to assess data quality in conjunction with of the evaluation of the quality of the indicators. The methodology was explained to the USAID staff at a meeting held at their offices in Pretoria on the morning of August 4, 2003 where the Team addressed issues and concerns pertaining to the approach.

The methodological approach was dictated to some degree by the list of partners originally supplied to the Team being an unknown percentage of the total partner Universe reporting against the SO6 and GCC indicators. The decision was therefore taken, and agreed to by USAID/SA, that the DQA would be based on Probability-Based Sampling of the given population, understanding that inherent limitations exist with sampling a pre-existing sample. The given list of partners was subdivided into three Populations - those that report data directly to USAID/SA, those that report data through a contractor and those for which data will be reviewed at USAID/SA.

Three (3) of the partners on the original list given to the Team, were purposefully excluded from the sampling for the DQA. These partners are DeLoitte and Touche; Deloitte and Touche Emerging Markets Group; and Agrilinks (EM&I). These partners were part of the SO5 DQA.

The Team's approach to assessing the quality of data collected and reported on by the individual partners was based on the internationally recognized International Organization for Standardization (ISO19011) "Guidelines for Quality and/or Environmental Management Systems Auditing" (Appendix A). This involved a standard data verification process on site that was administered by the Team. The approach required that partners complete the Data Quality Assessment Checklist (Appendix B) prior to the on-site visit. All partners received the checklist electronically and were notified of the requirement to complete it. The Team then reviewed the information presented in the checklist and performed the verification process of the data in accordance with the ISO 19011 guidelines[2]. The results of the validation process allowed the Team to assess each organization's capacity to collect and report on SO5 indicator data, and to point out strengths and vulnerabilities in the partners' data systems. This information is contained in summary form in sections 5, 6 and 7 of the report titled, "Data Quality Assessment."

The indicators were also assessed for their quality characteristics using a standardized assessment tool (Appendix C: Indicator Quality Worksheet). Each indicator was assessed for directness, objectivity, practicality and adequacy. Only the SO6 indicators were included in this detailed assessment.

1.6. Work Plan

Attached is the work plan and calendar (Appendix D) containing the key benchmarks and corresponding Team responsibilities for this exercise. In brief the activities performed for this DQA were as follows:

c. Preliminary review of the DQA documentation (ADS guidelines, TIPS, PWC Toolkit, etc), and preliminary discussion with Mega-Tech to identify priority issues & concerns.

d. Initial consultations with mission personnel to:
- Review the overall scope of work for the SO6 DQA,
- Clarify the set of indicators to be covered; and
- Discuss mission and bureau issues/concerns about indicators and data quality.

c. Preparation of performance indicator and data quality assessment tools to:
- Guide Team about indicator and data quality issues, and
- Provide the information source for indicator and data quality assessment tables.

[2] See section 3.3 for methodology.

d. Intensive consultations with implementing partners, where possible, to:
   - Conduct a validation exercise of data quality with each partner based on the information contained in Data Quality Assessment Checklist using the ISO 19011 audit guidelines method, and
   - Identify data strengths, vulnerabilities and risks.

e. Quality assurance and quality control reviews of incoming data validations.

f. Drafting of the various sections of the report in accordance with the time frame contained in the attached work plan calendar.

1.7. Detailed Data Quality Assessment Methodology

The DQA was based on the comparison of the audit evidence provided by the individual partners with the quality criteria for data, as set out in Appendix B. These were validity, reliability, timeliness, precision and integrity. In essence the purpose of the DQA audit was to establish whether there are any significant areas of strength or concern in each of the partners' ability to manage data to the highest level of validity and accuracy. In part, because the great variation in the nature of the partners' activities and operations, the audit was based on the definitions used by the partners themselves. This formed the most appropriate method to test their data quality management systems.

The audit technique was based on a sampling of evidence, which was required to be both valid and verifiable, to determine whether the partner met or was able to meet the set quality criteria. As a sampling technique was used, as is standard audit practice, it is not possible to confirm with 100% accuracy whether the partner meets all the criteria, in every circumstance, and thus the audit has some inherent limitations. Multiple techniques were used during audit to gather and verify information including observation, interview, document review and data review.

Audit findings are the results of the evaluation of the collected audit evidence against the audit criteria and were defined as follows:

a. A **NON-CONFORMITY** was declared when the audit evidence showed that there had been non-fulfillment of a criterion. Such non-conformities were classified as **MINOR** or **MAJOR depending the risk that the non-conformity presents to USAID/SA**.
   - A **minor non-conformity** indicated a failure to meet a required data quality criterion. Despite this failure the overall data quality characteristic e.g. validity, could still be achieved and only minimal risk is presented to USAID/SA.
   - A **major non-conformity** indicated a failure to meet a required data quality criterion. This failure prevented the achievement of the overall data quality characteristic and/or presents a significant risk to USAID/SA.

b. An **OBSERVATION** was noted when a **STRENGTH** or **VULNERABILITY** was noted in the partner's data quality system, which importantly, was not a non-conformity.
   - **Strengths** are identified with the purpose of giving positive feedback, which allows for a partner to focus on those areas of operation, which may be less effective and efficient.
   - **Vulnerabilities** are identified with the express purpose of giving the partner information on areas, which if not managed, may in the future result in a criterion not being fulfilled. They are an indication of risk and extremely useful for internal management use.

c. Achievement of the criteria by the partners was scored in order to give an overall assessment of the partners' conformance to the data quality requirements. The scoring rubric was as follows:
   - Three (3): Indicating that the partner met the criterion in its stated form.
   - Two (2): indicating that a minor non-conformity had been identified.
   - One (1): Indicating that a major non-conformity had been identified.

   The results of the DQA, per quality attribute, are averaged over the number of required criterion. The results are presented as a nominal scale between one (1) and three (3) where one (1) indicates an absolute failure/risk to achieve the required attribute and three (3) indicates that ideal data quality is achieved.

4. Data Quality Assessment Population One

1.8. Introduction

The results and recommendations associated with the Data Quality Assessment (DQA) verification audits for each of the partner organizations, per performance indicator, are given below. By assessing the data quality associated with each indicator, per partner, this DQA has evaluated the ability of each of the partners to meet the USAID basic data quality requirements.

1.9. Summation of DQA Results for Population One

A total of nine partners were reviewed in population one. Of these partners, one audit was declared not feasible due to the absence of any auditable information on the day of audit. For each partner the results of the DQA, per quality attribute, are presented as a nominal scale between one (1) and three (3) where one (1) indicates an absolute failure to achieve the required quality attribute and three (3) indicates that ideal data quality is achieved[3].

The average scores per quality characteristic are presented in Figure 1. This clearly demonstrates that the risk presented to USAID in terms of data quality is least in terms of the integrity of the data, i.e. it can be demonstrated that the data is free of undue or unethical manipulation. Greatest risk is presented in terms of the precision of data where in many cases there are high levels of error inherent in the data. This area also shows the greatest variability amongst partners with scores ranging from the minimum of 1 to the maximum of 3. The low scores attained by some partners with regards to data validity occur when data is 'tortured' in an effort to make it match the indicator reported on. This is a practice, which must be actively discouraged by USAID/SA.

Where reliability is low this is usually due to variations in collection methodologies and the difficulty of achieving repeatable results. Low scores for timeliness are most often the result of the reporting of data being inadequate in terms of allowing USAID to identify or manage any potential data-related risk.

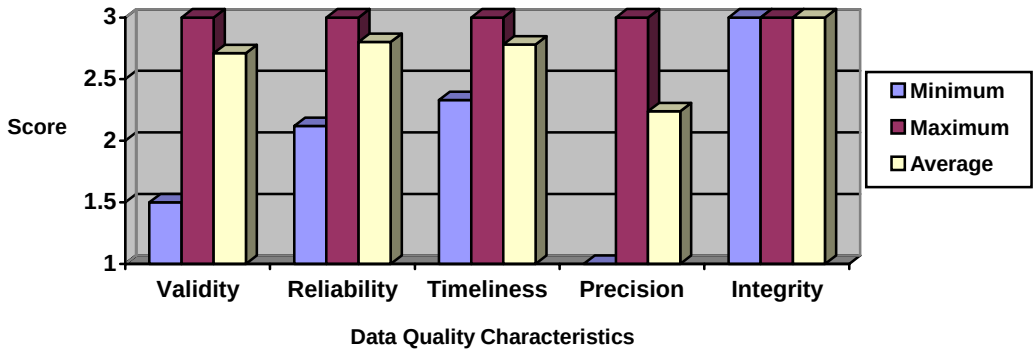

Figure 1: Average scores per quality characteristic for population one.

[3] See methodology for details.

1.10. Ndlandlamuka Local Project

Risk to USAID/SA:

As this agreement has been closed out, no prospective data risk is presented to USAID/SA in this regard. The risks are limited to the retrospective vulnerabilities that have been raised (Appendix E).

The Agreement:

NLP provided the Team with a copy of the agreement and the amendments. The NLP agreement with Mega-Tech was effected from 14 June 2002 and the expected completion date was 14 June 2003. The grant was originally for R 870 000 and the GMAC commitment was USD 87 000. The agreement, including some of the milestones, was amended on 02 December 2002 and again on 06 February 2003. The overall grant decreased to an eventual R 783 000. The project has already been completed and the close-out payment made to NLP. This partner should perhaps have been included in population two as it reported through MT. However it was included in the work-plan under population one and accepted as such prior to the physical audits.

Deliverables:

According to the Grant Agreement, NLP would report on "one USAID SO6 indicator semi-annually in the format to be provided by GMAC". No indication is made as to which indicator this actually is. The agreement describes six (6) overall tasks, comprising of in total 14 milestones, which needed to be completed and reported on the completion of each task. After the amendments only five (5) overall tasks remained with a total of 12 milestones. The milestones of interest to this DQA are:

- 100 beneficiaries trained.
- 300 trees planted and maintenance agreement signed.

The auditee was not sure which specific indicators NLP reported against. The NLP reports are based on the milestones listed with each task. Four (4) food gardens were initiated, two directly at two schools; one was to be launched at the end of August 2003; and a fourth that had to be stopped due to water problems. Receipts for items bought, e.g. seedlings, fertilizer, etc., were seen (e.g. Hydespray (Pty) Ltd Receipt No. 28919 dated 18/06/2003). No nurseries could be started due to the decrease in the grant value. 1211 trees were distributed instead of only 300 trees. A peanut butter plant was initiated and aided financially.

Reporting Validity:

According to the table of partners given the Team, NLP reported against GCC 2.1 only but on scrutinizing the GCC data tables provided by Mega-Tech, NLP reports on GCC 1.2, 2.1 and 2.4.

On comparison of the data reported and the indicators, the validity of reporting against GCC 2.1 for the number of trees planted is demonstrated. A minor discrepancy exists with regards the number of trees reported in the GCC data tables obtained from Mega-Tech – 1 235 trees are reported here when NLP reported only 1 211 trees on their Grant Activity Completion Report. Also only 1 211 trees could be verified at audit as being bought from Madken Nursery by NLP. Original letters of request for trees are kept together in a file along with the Tree Maintenance Agreement Agreements. The auditee pointed out that not all these agreements were in the file due to the financial audit in progress during the DQA. The audit does not demonstrate the validity of NLP reporting against GCC 2.4 – there was no evidence at audit nor in the reports from NLP that referred to leveraged funds even though there is a reference to leveraged funds from the Nelson Mandela Children's Fund against GCC 2.4 in the GCC data tables obtained from Mega-Tech. GCC 1.2 refers to capacity building that is relevant to milestones set in the NLP agreement – the audit demonstrates that reporting against this indicator is valid.

Previous DQA's:

Although NLP stated that USAID/SA and Mega-Tech representatives visited NLP on a few occasions, no records of such visits were made available to the Team. It must be noted here that Mr. Steve Horn of Mega-Tech admitted verbally to the Team that no record of the visits Mega-Tech made to partners were recorded in writing.

Limitation(s) to DQA:

A financial audit was in process on the date of the on-site visit and thus not all documentation was available.

Current DQA's:

The details of the DQA results for the two indicators on which NLP reported are given below. NLP filled in the DQA checklist before the on-site audit. The NLP answers/comments are entered into the DQA for interest. Where no answer/comment is shown the DQA checklist was left blank. The indicators are:

- GCC Indicator 2.1 Area where USAID has initiated interventions to maintain or increase carbon stocks or reduce their rate of loss.
- GCC Indicator 1.2 Increased capacity to meet requirements of the UNFCCC, including activities in land use/forestry and energy/industrial/ urban sectors.

1.10.1. NLP DQA1 of 2

Result:	GCC Result 2 – Reduced net greenhouse gas emissions from the land use / forest management sector									
Indicator:	GCC Indicator 2.1 – Area where USAID has initiated interventions to maintain or increase carbon stocks or reduce their rate of loss									
Score:	**V**	2.87	**R**	2.86	**T**	2.5	**P**	2.33	**I**	3.00
Data Limitations (Non-conformities):										
Non-Conformity 1 Data Manipulation: **MINOR –** There is missing data in the form of the records used to verify the number of trees actually planted at the locations visited. "Memorandum of Goods or Services Required" sheets are evidence that these visits occurred but no record of the visit itself was available – these documents have not been obtained from the person that made the location visits to count the trees actually planted. ***Non-Conformity 2 Representativeness of Data:*** **MINOR** – Not all units of the population had an equal chance of being selected for the sample – visits to locations to count the trees actually planted. Those places where less than 10 trees were donated were not considered for a location visit. ***Non-Conformity 3 Transparency:*** **MINOR** – Data quality problems are not mentioned at all in the Grant Activity Completion Report. ***Non-Conformity 4 Frequency:*** **MINOR** – This project was only for one year and data was only reported once at the end of the grant period. This may result in a possible risk as should there have been problems wrt data quality then it would have been too late for USAID/SA to implement corrective actions related to data quality issues. ***Non-Conformity 5 Precision:*** **MINOR** – the margin of error has not been determined however only a simple calculation is required as all the systems are there.										
Strengths and Vulnerabilities:										

Strength 1 Validity:

DM correlated the number of trees bought from Makden Nursery with the total number of trees allocated on the tree maintenance agreements. NLP also made on-site visits of all those recipients that were given more than ten (10) trees – and the trees physically counted. This did produce a problem in that not all could be visited due to resources and the distances involved.

Strength 2 Precision:

The pure simplicity of the system makes the margin of error easily calculated.

Strength 3 Integrity:

There is little room for inappropriate manipulation of data.

Vulnerability 1 Currency:

Although the "grant period" is sufficient a time period if referred back to the grant agreement, it would be better to give the actual time period the report refers to, to avoid ambiguity.

Vulnerability 2 Integrity:

No written records on previous visits to NLP were seen at audit.

Recommendations:

Recommendation 1 Validity:

There must be an audit trail to demonstrate the completion of the deliverable, i.e. from the acquired trees to the follow up of the planted trees. There is a gap between the acquisition of trees and the verification of the actual planted trees. The intent is there because there is a process in place to count the actual planted trees but NLP must make sure that they have access to the verification documents. The missing data must be obtained from the person that did the on-site location visits.

Recommendation 2 Transparency:

Encourage the partners to report on any issues that may impact on data quality.

1.10.2. NLP DQA 2 of 2

Result:	GCC Result 1 - Increased Participation in the UNFCCC									
Indicator:	GCC Indicator 1.2 - Increased capacity to meet requirements of the UNFCCC, including activities in land use/forestry and energy/industrial/ urban sectors									
Score:	**V**	3.00	**R**	2.86	**T**	3.00	**P**	2.33	**I**	3.00

Data Limitations (Non-conformities):

Non-Conformity 1 Transparency:

MINOR – Data quality problems are not mentioned at all in the Grant Activity Completion Report.

Non-Conformity 2 Precision:

MINOR – The margin of error has not been determined.

Strengths and Vulnerabilities:

Strength 1 Integrity:

There is little room for inappropriate manipulation of data.

Vulnerability 1 Face Validity:

The partner reports number of people capacitated whereas the GCC data tables obtained from Mega-Tech is reported as training activities.

Vulnerability 2 Face Validity:

ONLY if attending a training activity can be equated to "increased capacity", can there be a logical relationship between the activity and the indicator. However "increased capacity" refers to the actual increased understanding and way of thinking wrt to the training activity and this has not been measured.

Vulnerability 3 Frequency:

This project was only for one year. This may result in a possible risk because should there be problems wrt data quality then it would be too late for USAID/SA to implement corrective actions for "correct" data collection.

Vulnerability 4 Currency:

Although the 'grant period' is sufficient a time period if referred back to the grant agreement, it would be better to give the actual time period the report refers to, to avoid ambiguity.

Vulnerability 5 Integrity:

No written records on previous visits to NLP were seen at audit.

Recommendations:

Recommendation 1 Reliability:

Encourage the partners to report on any issues that may impact on data quality.

1.11. Chemonics International – Retail Water Distribution Project

Risk to USAID/SA:

The risks presented by RWD do not have their origins at RWD – they are mostly systems issues such as the management of indicators. RWD have tried to determine data on something that is almost immeasurable – sustainability. RWD will apparently be correcting their sampling method for the annual 2003 report – USAID/SA should ensure that this is verified before accepting data from RWD otherwise the same data quality risks will be repeated (Appendix F).

The Agreement:

The RWD agreement with USAID/SA was effected from 21 September 2000 and the expected completion date is 30 May 2004. The grant is for a CPFF of USD 2 299 950 with an obligatory amount of USD 1 000 000. The role of RWD is to assist local government to develop policies and procedures for their new role as a Water Support Authority (WSA) and to assist with the implementation of these decisions.

Deliverables:

According to the agreement (page 13), RWD would report on the "number of low-income communities in which sustainable environmental practices are being applied". This indicator is not on the PMP (03 August 2003 version) that was given the Team but it is one of the indicators under Intermediate Result 6.4 according to the description in the agreement. RWD has no idea of the indicator(s) he is reporting against and reported against PI 6.4.1 with his last annual report based on documentation sent to RWD from USAID/SA.

RWD considers water services to be a standpipe with 100% running water 200m distances from the home or in-house 100% running water. Persons were trained to maintain the systems. Reasons for interrupted water service:

- Electricity account not paid for electric boreholes so Eskom cuts off electricity.
- No money to buy diesel for diesel boreholes.
- Drought – water level low therefore no flow into pipelines.

Reporting Validity:

According to the table of partners given the Team, RWD reports against SO 6.2(a) and GCC 4.1 but on scrutinizing the PMP and the GCC data tables provided by Mega-Tech, RWD reports on SO 6.1(a); SO 6.2(a) PI 6.2.2 and GCC 4.1. On comparison of the data reported and the indicators, the Team agrees with the validity of reporting against SO 6.1(a) and SO 6.2(a). RWD are ONLY involved in providing a municipal service of drinking water – so reporting money leveraged for drinking water and number of households receiving water is valid. Although PI 6.2.1 (Rand value) and PI 6.2.2 (number of households) are for municipal services in total and thus includes water services, the Team finds reporting against these indicators not valid. It is clear that aggregation of data from different sources and different data types is involved, placing the validity of that which is eventually reported in question. RWD also report Level of Effort (LOE) on a quarterly basis – these are any energy conservation activities related directly to water services provision. This was not audited at time of audit. GCC 4.1 was not audited at time of audit. GCC 4.1 appears to be a list drawn up by USAID/SA and/or Mega-Tech.

Previous DQA's:

Although RWD should be reporting on IR 6.4, USAID/SA (Sergio Guzman) visited RWD on 04 November 2003 and reported on a DQA completed against PI 6.2.1 and PI 6.2.2. At audit, no evidence of the 13 000 households (PI 6.2.2) mentioned in this report was found. RWD reported number of communities in the previous annual report (facsimile dated 24 October 2002) and also stated that they intend reporting number of communities again this year.

Limitation(s) to DQA:

The performance indicator PI 6.4.1 was not in the PMP (dated 03 August 2003 version) given the Team. Upon investigation PI 6.4.1 was sent to the Team but from the May 2003 version. The DQA was completed against SO 6.2(a). A second DQA was not completed against PI 6.4.1 because the Team had a fundamental problem with this indicator – the definitions of community and sustainability. Two documents required to verify two Rand values reported on "Attachment 9" in the previous annual report, were not on the RWD premises. These documents are kept at the Metro offices.

Current DQA's:

The details of the DQA result for two indicators on which RWD report is given below. These indicators are:

- SO 6.1(a) Rand value of municipal services – water
- SO 6.2(a) Number of households receiving municipal services – water.

The next annual report – Oct 2003:

The following have been initiated and in some instances completed by RWD to improve the manipulation of the data –

- Aerial photographs of May 2003 will be used to count the number of houses in each community and to calculate the actual population based on the government definition of 1 household = 6 people.
- This will be correlated with numbers obtained from tribal authorities for population as well as boreholes and pipelines.
- A separate account for water services has been opened for the municipalities.

Due to the investigation completed on all the communities, RWD will be able to report on ALL the communities for the next annual report in Oct 2003. RWD is hoping to do this investigation on an annual basis.

1.11.1. RWD DQA 1 of 2

Result:	IR 6.4 Improved capacity to apply sustainable participating environmental management principles to local level development									
Indicator:	SO 6.2(a) Number of households receiving municipal services – water									
Score:	**V**	1.50	**R**	2.89	**T**	2.33	**P**	1.60	**I**	3.00
Data Limitations (Non-conformities):										

Non-Conformity 1 Face Validity:

MAJOR – Although water services is measured in both the indicator and by RWD, the absolute risk is introduced by the fact that the indicator measures number of households but RWD reports number of communities – the numbers will be completely different to what is expected by USAID/SA.

Non-Conformity 2 Sampling Error:

MAJOR – RWD could not show the questions asked at the oral survey so the survey questions could not be audited.

Non-Conformity 3 Non Sampling Error:

MAJOR – RM has an elaborate spreadsheet in an attempt to determine sustainability and actual population. The definitions are not operationally precise.

Non-Conformity 4 Data Manipulation:

MAJOR – RM verbally described an elaborate 'extrapolation' as to how the eventual 4 communities were reported. This could not be audited at all because neither the survey questions nor the calculation used could be shown for audit – this introduces an absolute risk.

Non-Conformity 5 Representativeness of Data:

MAJOR – Only 5 of a total of 135 communities were used for reporting purposes – only 3.7% of the total population. This is an absolute risk. Besides the small population percentage, only the proclaimed towns (the 5 communities) were considered for reporting purposes – the other communities are all rural (previous R293) villages.

Non-Conformity 6 Transparency:

MINOR – Data quality problems are not mentioned in the report.

Non-Conformity 7 Frequency:

MINOR – The data has only been reported once so far and reporting is only done annually. This may not be frequent enough to inform program management decisions.

Non-Conformity 8 Currency:

MINOR – No date of collection is identified in the report.

Non-Conformity 9 Precision:

MINOR – The margin of error has not been determined.

Non-Conformity 10 Data Source Type:

MINOR – The possible risks associated with municipalities as data sources have has not been identified.

Strengths and Vulnerabilities:

Vulnerability 1 Consistency:

Only one report has been handed in so the consistency of the sampling method could not be audited. USAID/SA should be aware that should the same sampling method be used in the 2003 report then there will be the same risks highlighted in this DQA. It must be noted that RWD have changed all their processes and will apparently be reporting on all the communities for 2003.

Vulnerability 2 Transparency:

RM was not able to show the e-mail records to the supervisor without having to do a major search in his 'Inbox" – the records were not insisted on.

Recommendations:

Recommendation 1 Validity:

Encourage the partners to ensure they have a verifiable audit trail for all data they report to USAID/SA.

Recommendation 2 Reliability:

Encourage the partners to report on any issues that may impact on data quality.

1.11.2. RWD DQA 2 of 2

Indicator:	SO 6.1(a) - Rand value of municipal services – water									
Score:	**V**	2.93	**R**	2.67	**T**	2.40	**P**	2.33	**I**	3.00
Data Limitations (Non-conformities):										
Non-Conformity 1 Face Validity: **MINOR** – RWD is reporting total leveraged funds for the project as a whole not for water services only, which is what the indicator description requires. ***Non-Conformity 2 Internal Quality Control:*** **MINOR** – There are no procedures for periodic review of data collection or maintenance. ***Non-Conformity 3 Transparency:*** **MINOR** – Data quality problems are not mentioned in the report. ***Non-Conformity 4 Timeliness:*** **MINOR** – The data has only been reported once so far and reporting is only done annually. This may not be frequent enough to inform program management decisions. ***Non-Conformity 5 Precision:*** **MINOR** – The margin of error has not been determined.										
Strengths and Vulnerabilities:										
Nil										
Recommendations:										
Recommendation 1 Timeliness: Encourage the partners to report on a more frequent basis – not just annually.										

1.12. Isandla Partners in Development

Risk to USAID/SA:

Due to the well-established and consistent data systems at IPD, no present risks with respect to data quality were identified during this DQA (Appendix G).

The Agreement:

This IPD agreement with USAID/SA was effective from 01 June 2001 and the close out date was February 2003. The grant was for USD 150 000 from USAID with USD 50 000 contribution from another counterpart. The full USD 150 000 was obligatory. IPD are involved in the entire housing and services process including the actual implementation of the house building. The agreement has been closed out (February 2003) and IPD have received their final payout.

Deliverables:

According to page 8 of the agreement, IPD would deliver "… 160 new houses that incorporate energy saving features …" (amongst other deliverables). WRT indicators IPD had to report on the following indicators (page 9 of the agreement): "…

- Number of historically disadvantaged households assisted to obtain shelter and services.
- Amount of credit or subsidies leveraged.
- Emission of carbon dioxide equivalents avoided.

IPD will provide baseline data, targets and actual results. Reports on these indicator will be submitted very 6 months. …". This wording is similar to that of PI 6.2.2; PI 6.2.1 and PI 6.4.3 respectively.

IPD was unsure as to which project was to be audited so had not really prepared anything as such – however most requested documentation was close at hand. IPD were uncertain of the indicators they were reporting against and in fact had the wrong IR definitions on the table they used for reporting purposes. IPD reported using a tabled spreadsheet designed by them. This table includes a column for the data source used to obtain the data reported.

Reporting Validity:

According to the table of partners given the Team, IPD reports against PI 6.1.3; PI 6.2.1; PI 6.2.2; and no GCC indicators but on scrutinizing the PMP and the GCC data tables provided by Mega-Tech, IPD reports on SO 6.1(e); SO 6.2(e); PI 6.1.3; PI 6.2.1; PI 6.2.2; and GCC 1.2. IPD apply for separate subsidy approvals by ID number and erf number. The subsidy approval from PHB is used to obtain data for both number of households and Rand value because the subsidy for all the houses are exactly the same Rand value.

Previous DQA's:

USAID/SA did a DQA on 06 January 2003 on SO 6.1(e); SO 6.2(e); PI 6.1.3; PI 6.2.1 and PI 6.2.2.

Limitation(s) to DQA:

The computer of the Project Manager, Liesel du Plessis, "crashed" recently resulting in that she could not explain some of the calculations on the spreadsheet – she could not remember the calculations used.

Current DQA:

The detail of the DQA results for two indicators on which IPD report is given below. Only one DQA has been completed because the same data source is used to obtain data for both indicators. These indicators are:

- PI 6.2.1 Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision.
- PI 6.2.2 Number of households assisted to obtain shelter/urban services through the provision of credit and subsidies to low-income communities.

1.12.1. IPD DQA 1 of 1

Result:	IR 6.2 - Previously ineligible households developers, builders and municipal service providers obtaining access to credit									
Indicator:	IR 6.2.1 - Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision. **AND** IR 6.2.2 - Number of households assisted to obtain shelter/urban services through the provision of credit and subsidies to low income communities.									
Score:	**V**	2.93	**R**	3.00	**T**	3.00	**P**	2.33	**I**	3.00
Data Limitations (Non-conformities):										
Non-Conformity 1 ***Data Manipulation:*** **MINOR** – Some aggregations did not total correctly on the report. ***Non-Conformity 2*** ***Precision:*** **MINOR** – The margin of error has not been determined.										
Strengths and Vulnerabilities:										
Nil										
Recommendations:										

Nil

1.13. Peoples Housing Partnership Trust

Risk to USAID/SA

There is little audit evidence to suggest that the data quality practices of this partner present any significant risk to USAID/SA (Appendix H).

Agreement and deliverables:

The PHPT provides technical assistance and policy support to "The People's Housing Process" a housing delivery system, which is classified as one of the major contributors to the subsidized housing stock. The PHPT is a parastatal situated within the Department of Housing in Pretoria, and operates as if it were essentially part of it. It operates through the Provinces, but also provides technical assistance to municipalities and community groups.

Reporting Validity:

The national housing data base distinguishes between PHP units and all others. This therefore provides an easily auditable source of data. The PHPT claims that all housing developed through the PHP can be attributed to them. Since USAID is a major funder of the PHPT, the total of the annual PHPT output is attributed to USAID. PHPT has no difficulty with this. However, this question of attribution is an important one, especially noting the very large numbers that are involved.

1.13.1. PHPT DQA 1 of 2

Result:	IR 6.1.3: Number of shelter units completed								
Indicator:	6.2 (e) Number of households receiving services – housing								
Score:	**V**	3.00	**R**	2.89	**T**	3.00	**P**	3.00	**I**	3.00

Data Limitations (Non-conformities):

Non-conformity 1: ***Transparency***

MINOR - Data collection, cleaning, analysis, reporting, and quality assessment procedures are not documented in writing

Strengths and Vulnerabilities:

Strength 1: ***Validity***

The unit of measurement is well defined and represents the focus of the program

Strength 2: ***Validity***

The data is obtained from well audited primary sources

Strength 3: ***Reliability***

The data are part of a national, well-audited data base which has been in use for ten years.

Strength 4: ***Timeliness***

The data is produced from a live data base managed by each Province, and can be accessed at any time

Strength 5: ***Precision***

No sampling is involved

Strength 6: ***Integrity***
Data is cross checked against auditable facts
Strength 7: ***Data Source Type***
Well tested and audited data is used
Vulnerability 1: ***Validity***
Attribution of all housing developed through the "People's Housing Process" to the PHPT, and by inference to USAID raises some difficulties, especially in light of the very large numbers reported
Vulnerability 2: ***Reliability***
The data cleaning and checking process is undertaken by a third party – the Provinces – and cannot be verified.
Recommendations:
Nil

1.13.2. PHPT DQA 2 of 2

Result:	6.2.1 Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision									
Indicator:	6.1 (e) Rand Value of municipal services completed - housing									
Score:	**V**	2.62	**R**	3.00	**T**	3.00	**P**	2.40	**I**	3.00
Data Limitations (Non-conformities):										
Non-conformity 1: ***Validity***										
MINOR - The data for reporting on this is the number of units: the partner does not report on Rand Value. In practice the range of subsidies provided is very small, so this vulnerability is minor										
Non-conformity 2: ***Precision***										
MINOR – Although no margin of error has been established, the likely margin is less than the change been affected by the project.										
Strengths and Vulnerabilities:										
Strength 1: ***Reliability***										
The primary data collection has followed a consistent and easily auditable process										
Strength 2: ***Timeliness***										
Data are taken from a live data base										
Vulnerability 2: ***Precision***										
The Rand value is based on a standardized subsidy amount, not on actual expenditure.										
Vulnerability 3: ***Data Source Type***										
The risk of using standardized proxy data for the subsidies instead of actual amounts has not been established.										
Recommendations:										
Recommendation 1: ***Validity***										
The partner should be requested to report the Rand Value										

1.14. Municipal Infrastructure Investment Unit

Risk to USAID/SA:

It is very difficult to estimate with absolute accuracy the Rand values of long-term agreements lasting more than 15 years. MIIU are not in a position to report actual Rand values and therefore report estimated nominal Rand values of these agreements in their entirety. This presents an absolute risk, in that the margin of error cannot be calculated. Reporting of these estimated nominal Rand values may have inherent errors and the margin of error may increase with each calculation to determine the estimation (Appendix I).

The Agreement:

This MIIU agreement with USAID/SA is a "performance-based completion-type Cost Plus Fixed Fee (CPFF)" agreement and was effective from 28 December 2001. The agreement has an estimated CPFF value of USD 2 255 025 with an obligation of USD 600 000. The role of MIIU is purely to bring various parties together for implementation of projects ("linkage officers") and to source funds. MIIU have no role in the actual building of houses and mostly have no access to the project once implementation starts. It is therefore difficult for MIIU to obtain data, e.g. Rand values of the actual monies spent and the number of households serviced, on any project.

Deliverables:

According to page 12 of the agreement, MIIU would report on "… these measure the Rands leveraged and the households served under programs supported by the MIIU. These can be examined through USAID/SA's website …".

MIIU have no idea which indicator(s) they are reporting against nor do they know what USAID/SA does with the data MIIU reports to USAID/SA. MIIU presently report using a tabled spreadsheet with a record of all the projects they have been involved in since 1998.

Reporting Validity:

According to the table of partners given the Team, MIIU reports against SO 6.1(a-d); SO 6.2(a-d); PI 6.1.3; PI 6.2.1; PI 6.2.2 and GCC 3.5 but on scrutinizing the PMP and the GCC data tables provided by Mega-Tech, MIIU reports on SO 6.1(a-d); SO 6.2(a-d); PI 6.1.3; PI 6.2.1; PI 6.2.2 and GCC 1.2; 3.1 and 3.5.

MIIU report estimated nominal Rand values. Even the 'Total Households in Municipal Area' is estimated from municipal records and is by no means accurate especially for holiday areas such as Overstrand. It is therefore not valid for MIIU to report number of households. MIIU can report estimated nominal Rand values for USAID/SA to obtain an idea of the size of the projects handled by MIIU however USAID/SA must bear in mind the risk involved in reporting these estimated nominal Rand values. It is therefore not valid for MIIU to report against indicators SO 6.2(a-d) PI 6.1.3 and PI 6.2.2.

Also because MIIU report estimated nominal Rand values, segregation of municipal services is not practical. Involvement from MIIU basically stops after the agreement between the parties is signed. Rand values are reported 'before the fact' - the actual cost cannot be reported – a total for municipal services is reported. In fact, even the projects that are dealing with separate municipal services have problems wrt reporting. An example of this is the 'water' sector projects – these include the necessary pipelines for both drinking water and sewerage. MIIU averaged the ratio to 67/23:drinking water/sewerage from three completed projects – Dolphin Coast, Nelspruit and Harrismith. MIIU uses this ratio to report these services separately when requested. It is therefore not valid for MIIU to report against indicators SO 6.1(a-d).

Previous DQA's:

USAID/SA did a DQA on 08 May 2002 only on IR 6.2. In this report (dated 10 May 2002), it is stated "… The MIIU calculates the number of households served based on the serviced area. In most projects, (Kelvin Power Plant for example) they use the actual number of accounts held by the utility. These are counted as individual households. For other services such as water, they estimate the serviced area and use demographic information provided by the municipality and/or utility to estimate the households served unless individual account information is available. During this assessment, the data that MIIU has reported for FY2001 was reviewed and it was found that for both the number of households and services data, the data was considered to be reasonably valid ...". This confirms most of what is mentioned under the Reporting Validity heading and

yet the data was still accepted without question. An audit is either valid or it is not – there are no 'reasonability-levels'.

Limitation(s) to Current DQA:

- Many of the agreements MIIU facilitates are long term (more than 15 years). Because values change over such a long period, actual Rand values are not even mentioned in these agreements. MIIU asks various independent consultants to estimate the value of each project; they then use these calculated estimated values to report to USAID/SA.
- DQA was reviewed at MIIU level only and not with the independent consultants.
- MIIU was in the process of moving to their new premises at the time of audit and most of their files were still in boxes, however most of the documentation was available.

Current DQA:

Based on the explanation above it was decided to only do one DQA on Rand value. It is still unclear due to the nature of the projects exactly under which indicator(s) these values would fall but the closest to the given indicators is PI 6.2.1. The detail of the DQA results for one indicator on which MIIU report are given below. This indicator is:

- PI 6.2.1 Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision

1.14.1. MIIU DQA 1 of 1

Result:	IR 6.2 - Previously ineligible households, developers, builders and municipal service providers obtaining access to credit									
Indicator:	PI 6.2.1 - Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision (leveraged funds by definition)									
Score:	**V**	2.85	**R**	2.12	**T**	2.80	**P**	1.00	**I**	3.00

Data Limitations (Non-conformities):

Non-Conformity 1 Non Sampling Error:

MAJOR – Although it has been determined that USAID/SA would prefer partners to report actual Rand values, MIIU report estimated nominal Rand values. The indicator definition does not specify this distinction. The operational preciseness of the definitions cannot be determined, introducing an absolute risk to data being reported.

Non-Conformity 2 Consistency:

MAJOR – Although MIIU collect data as directly from the agreements or from letters/e-mails of independent consultants, the reliability of the data from these independent consultants is a risk – there are too many intricate calculations that introduce increasing margins of error with each calculation. This questions the reliability of these estimated Rand values.

Non-Conformity 3 Internal Quality Control:

MAJOR – MIIU cannot report actual leveraged Rand values because they have no contact with the parties of the agreement after signage. There is no way the reported values can be checked for accuracy.

Non-Conformity 4 Transparency:

MINOR – There is no periodic review of data collection, maintenance and processing. MIIU have no further contact with the parties of the agreements after signing unless a agreement is re-negotiated.

Non-Conformity 5 Transparency:

MAJOR – Data quality problems are not mentioned at all in the reports seen. Due to the unreliability of the estimated values this is a significant risk.

Non-Conformity 6 Currency:

MINOR – The period the data was collected cannot be determined exactly – giving only a year is not sufficient a time period. Important data can be excluded/included if the exact dates (at least months) are not reported.

Non-Conformity 7 Precision:

MAJOR – No margin of error has been determined and because MIIU does not obtain actual Rand value data, this cannot be measured in the present process leaving the system open to absolute risk.

Non-Conformity 8 Data Source Type:

MAJOR – The data provided by the tertiary data sources – independent consultants – use many intricate calculations that introduce increasing margins of error with each calculation. There is an inherent risk of reporting estimated values but values based on an intricate set of calculations is even more risky due to this error factor increase.

Strengths and Vulnerabilities:

Vulnerability 1 Currency:

All projects from inception in 1998 are reported over and over with each quarterly report. There is a possibility that these cumulative amounts are reported over and over to Washington thus reporting elevated incorrect results.

Recommendations:

Recommendation 1 Reliability:

Encourage the partners to report on any issues that may impact on data quality.

Recommendation 2 Currency:

Encourage partners to include the exact time period of their report on their report.

Recommendation 3 Precision:

MIIU should report actual Rand values unless USAID/SA are content and aware of the inherent risk involved in reporting estimated data.

1.15. Cato-Manor Development Association

Risk to USAID/SA:

Although no prospective data risk is presented to USAID/SA, the following significant risks must be noted as retrospective Non-Conformities:

MAJOR – There was no DQA evidence to show that data was reported on a grant/project that totaled USD 805 570.

MAJOR – The transcription error at USAID/SA of USD 6 million instead of R 6 million, places the validity of past reports in question.

The Agreement:

This CMDA agreement with USAID/SA was effective from 18 July 2001 and the close out date was 31 July 2003. The grant was for an obligated amount of USD 200 000 and counterpart funds to the value of USD 605 570. CMDA are involved in the entire housing and services process including the actual implementation of the house building. The agreement has been closed out and CMDA have drawn-down their final payment. The concept of this agreement changed completely from its original planning. The USAID monetary contributions for civils and top structure were reversed by an amendment, i.e. USD 32 911 was allocated to top structure (previously civils) and the USD 167 089 was allocated to civils (previously top structure).

Deliverables:

The project was handled completely differently from other projects – instead of building many houses of the same shape, show houses were built and these show houses are going to "sell" the houses over the next two years. Money was spent on the planning – the design process was looked at differently. Planners, engineers, etc. were brought together to find solutions for the difficult building conditions in KZN – the slopes, the quality of the soil, etc. House extensions were also considered. This would mean that not all the houses would look the same – the areas under consideration are in open view to tourists. There are 7 pieces of land of which only 2 areas have show houses. Once all the houses have been built landscaping will be initiated.

Reporting Validity:

CMDA have never reported data to USAID/SA. USAID/SA (Sergio Guzman) obtained the data that is reported in the GCC tables (3.5(b)) telephonically or via e-mail from CMDA. The USD 6 million reported in the GCC Table 3.5(b) should be R 6 million. On closure, the value of the project will be close to R 50 million.

Previous DQA's:

USAID/SA visited CMDA in the past (including a visit from Jimmy Carter) but no written records were sent to CMDA in this regard nor were the Team given any DQA's for CMDA.

Limitation(s) to DQA:

CMDA has been run as a Section 21 project but is in the process of closing down. CMDA are handing over everything to Ethekweni Metro and will no longer exist after September 2003. The entire filing system, including the project in question, had already been handed over to Metro. ("Double-auditing" is standard process at Metro and official financial audits against Generally Accepted Accounting Principles (GAAP) are completed on an annual basis). Metro will be handling the sale of the houses in the next phase of the project. The agreement could not just be amended as the partner has changed.

Current DQA:

Since CMDA have never reported data to USAID/SA, there was no data to audit on the day of the audit and the audit was therefore abandoned. Not even Rand values as spent on the design process (civils) have been reported.

1.16. Kwa-Zulu Natal Project Preparation Trust

Risk to USAID/SA:

The risks presented by PPTKN do not have their origins at PPTKN – they are systems issues such as the management of indicators. Due to the well-established and good systems at PPTKN, no other present risks with respect to data quality were noted at audit (Appendix J).

The Agreement:

PPTKN had a agreement with USAID/SA that was effective from 17 September 1993 and the close out date was 15 September 2001. The grant was originally for USD 400 000 but was amended on 12 September 1996 to USD 2 000 000. PPTKN reported against PI 6.2.1 and PI 6.2.2 for this agreement. PPTKN is presently in a brand new agreement. The new agreement is not with USAID/SA directly but with Mega-Tech. This new agreement started on 23 February 2003 and is planned to end on 27 September 2004. The grant is for R 960 000 and the GMAC commitment is USD 120 000.

Deliverables:

In the agreement with Mega-Tech, the standard, "report against at least one (1) SO6 indicator" is stated however PPTKN took it one step further when filling in the documentation from Mega-Tech. Page 4 of the GMAC Grant Application Form refers to SO 6.1(e) and PI 6.1.3. PPTKN are required to work on a minimum of 4 projects on this agreement but already have 11 projects that have been identified and initiated. PPTKN are 'linkage officers' and work on a slightly different manner to other partners – a 'bulk' subsidy approval is obtained. Once the 'bulk' subsidy is approved, they have no involvement in the actual implementation of the house building. PPTKN tries to keep their records updated with monies actually spent but find this sometimes difficult especially for the rural areas.

Reporting Validity:

According to the table of partners given the Team, PPTKN reports against PI 6.2.2 and on scrutinizing the PMP and the GCC data tables provided by Mega-Tech, PPTKN reports on PI 6.2.2. It is assumed that this was based on the first agreement. On comparison of the data reported and the indicators, it would not be valid to report against PI 6.2.1 and PI 6.2.2 – only subsidies are involved for this entire project. Although the descriptions of these indicators refer to "credits and subsidies", the Team has a fundamental problem with these two sources of monies being combined. SO 6.1(e) and SO 6.2(e) refer only to housing whereas total house plus infrastructure is being reported. On comparison of the data reported and the indicators, it would not be valid to report against PI 6.1.3. Rand value and number of households are reported before implementation of the project so completed shelter units would not be feasible. It was stated that disaggregation into the separate services would be possible but it must be remembered that it is based on 'bulk' approval values and that the predictive value is usually lower than actual eventual costs. PPTKN would like to see an integrated indicator, e.g. special needs such as HIV/AIDS also require housing and may presently be more important than economic development.

Previous DQA's:

USAID/SA and Mega-Tech have apparently visited PPTKN on this new agreement but no written evidence of this was seen at audit. There were written DQA's for the first closed out agreement.

Limitation(s) to DQA:

- All the information given the Team, was based on the first agreement that had already been closed out.
- Definitions of indicators.

Current DQA:

The same data source is used to obtain data to report both SO 6.1(e) and SO 6.2(e), i.e. only one DQA is required. The details of the DQA results for RWD are given below. These indicators are:

- SO 6.1(e) Rand value of municipal services – housing.
- SO 6.2(e) Number of households receiving services – housing.

1.16.1. PPTKN DQA 1 of 1

Indicator:	SO 6.1(e) - Rand value of municipal services – housing SO 6.2(e) - Number of households receiving services – housing									
Score:	**V**	2.85	**R**	3.00	**T**	3.00	**P**	2.33	**I**	3.00
Data Limitations (Non-conformities):										
Non-Conformity 1 Face Validity: **MINOR** – PPTKN are reporting directly from the bulk subsidy approval they receive from DoH. The agreement is to build x number of houses and the Rand value is determined by a simple multiplication as each subsidy for each house is exactly the same. The risk lies with the fact that the subsidy includes the infrastructure and is therefore not just for housing as is defined by the indicators. ***Non-Conformity 2 Precision:*** **MINOR** – the margin of error has not been determined.										
Strengths and Vulnerabilities:										
Strength 1 Overall: The simplicity of the data management system makes verification easy. ***Vulnerability 1 Integrity:*** No written records on visits to PPTKN were seen at audit.										
Recommendations:										

Recommendation 1	Ensure that written records are kept of all visits to all partners.

1.17. University of Cape Town

Risk to USAID/SA:

As this agreement has been closed out, no prospective data risk is presented to USAID/SA in this regard. The risks are limited to the retrospective vulnerabilities that have been raised (Appendix K).

The Agreement:

UCT provided the Team with a copy of the agreement and the amendment. This UCT agreement with Department of Environmental Affairs and Tourism (DEAT) and USAID/SA was effective from 23 June 1999 and the last report was handed in March 2003. The grant was for a 27 month period for USD 1 581 900 for the Awareness Component and R 2 183 000 for the Training Component. A levy amount of R143 809.09 for the Awareness Component was also awarded. At the end of the designated agreement period, there was still a substantial amount of the grant monies unused – the value of the grant had increased by R 550 000 due to the Rand/USD rate. An amendment was made to the agreement giving an extension on time as well as adding more deliverables –UCT worked as consultants for DEAT. The agreement has been closed out. No final closeout report has been requested from or given to USAID/SA as at the time of audit.

Deliverables:

The agreement basically only listed deliverables - no reporting against indicators as such. Reports were handed over to USAID/SA after each training activity and these reports included the evaluation of the training activity as well as the attendance list. The Awareness Component was aimed at politicians, councilors, senior officials and the like. The Training Component was aimed at capacity building of senior level government, provincial and local authorities and the like. UCT networked with other tertiary education facilities to do training in their respective areas/provinces. UCT designed an evaluation form that was distributed after each training activity. The results of these evaluation forms are given in each report. They also designed a second questionnaire that focuses on the measure of the effectiveness of the training activities. This questionnaire was filled in by UCT at focus groups, individual interviews or telephonic interviews 4 – 6 months after the training activity. They used this information for their own research purposes and will be issuing a paper on the subject when time allows.

Reporting Validity:

According to the table of partners given the Team, UCT reported against GCC 1.2 and on scrutinizing the PMP and the GCC data tables provided by Mega-Tech, UCT also only reports on GCC 1.2. On comparison of the data reported and the indicators, the Team agrees with the validity of reporting against GCC 1.2.

Previous DQA's:

No written record of DQA visits were present at the on-site audit, UCT in fact stated that although USAID/SA visited regularly, no written record has ever been forwarded to them for their records.

Limitation(s) to DQA:

- UCT reported to USAID/SA after each training activity giving a narrative wrt the evaluation of the training course. No other reports or a final closeout report has been handed in to USAID/SA with a summary of all the training activities run.
- No information was given the Team before the on-site visit – all information had to be obtained on-site.

Current DQA:

The details of the DQA results for one indicator on which UCT reported are given below. This indicator is:

- GCC 1.2 Increased capacity to meet requirements of the UNFCCC, including activities in land use/forestry and energy/industrial/urban sectors.

1.17.1. UCT DQA 1 of 1

Result:	GCC Result 1 – Increased Participation in the UNFCCC									
Indicator:	GCC 1.2 - Increased capacity to meet requirements of the UNFCCC, including activities in land use/forestry and energy/industrial/urban sectors									
Score:	**V**	2.67	**R**	2.33	**T**	3.00	**P**	2.00	**I**	3.00

Data Limitations (Non-conformities):

Non-Conformity 1 Transcription Error:

MINOR – There is no check on any capture of data and this has led to missing data as determined at audit. This has not been of concern until the time of the audit as they are more concerned with narrative – the evaluation part of the report.

Non-Conformity 2 Transparency:

MINOR – Data quality problems are not mentioned at all in any of the reports.

Non-Conformity 3 Precision:

MINOR – The margin of error has not been determined.

Strengths and Vulnerabilities:

Strength 1 Timeliness:

A report is written and sent to USAID/SA after each training activity.

Vulnerability 1 Face Validity:

ONLY if attending a training activity can be equated to "increased capacity", can there be a logical relationship between the activity and the indicator. However "increased capacity" refers to the actual increased understanding and way of thinking wrt to the training activity and this has not been measured.

Vulnerability 2 Transcription Error:

The emphasis in the reports was placed on the narrative whereas USAID/SA places emphasis on numbers.

Vulnerability 3 Transcription Error:

For reporting purposes, the number of attendees at training activities has been equated to "increased capacity". See Vulnerability 1 for explanation.

Vulnerability 4 Integrity:

Although DrMS stated that USAID/SA had visited UCT a few times, no written record was found on-site at audit.

Recommendations:

Recommendation 1 Transparency:

Encourage the partners to report on any issues that may impact on data quality.

1.18. Johannesburg Housing Company

Risk to USAID/SA:

Two problems were noted: the first was that in the reporting, units developed before the grant had been received were reported on. At about 500 units this makes a very small impact on the total numbers presented in the PMP. The second was that the financial year reported on by the Johannesburg Housing Company is July – June, but USAID collected data for the US Fiscal year. There was a potential misreporting for that reason but a check demonstrated that in neither of the two years would the numbers have been different if the US year had been used (Appendix L).

The Agreement and Deliverables

The Johannesburg Housing Company was given a grant to facilitate preparation of feasibility studies for housing in the inner city areas of Johannesburg. The funds were used to pay consultant fees. Some projects were found to be unfeasible, but most proceeded.

Reporting Validity:

The number of units attributed to USAID and included in the annual report were those that has been developed as a result of the feasibility studies undertaken. A house is only reported when it is complete and ready for occupation (i.e. a tenant has signed a lease). Thus attribution is direct and auditable.

1.18.1. JHC DQA 1 of 2

Result:	IR 6.2.2:Number of households assisted to obtain shelter/urban services through the provision of credit and subsidies to low income communities									
Indicator:	2 (e) Number of households receiving services – housing									
Score:	**V**	2.60	**R**	3.00	**T**	2.33	**P**	3.00	**I**	3.00
Data Limitations (Non-conformities):										
Non-conformity 1: ***Validity*** **MAJOR** - Data is aggregated from years prior to the grant agreement thus creating a major fault. ***Non-conformity 2:*** ***Timeliness*** **MINOR** - Data are available from a live database, but were only reported to USAID annually. Note: a quarterly report was produced but did not report on specific indicators.										
Strengths and Vulnerabilities:										
Strength 1: ***Validity*** The unit of measurement is well defined and represents the focus of the program ***Strength 2:*** ***Reliability*** The data come from a well-audited data base ***Strength 3:*** ***Precision*** No sampling is involved ***Strength 4:*** ***Integrity*** Data is cross checked against auditable facts ***Vulnerability 1:*** ***Timeliness*** The annual reporting period used is different from the USAID fiscal year.										

Recommendations:
Since the program is completed recommendations are not required

1.18.2. JHC DQA 2 of 2

Result:	IR 6.2.1 Rand value of credit and subsidies obtained from households for HDP shelter and service provision									
Indicator:	1 (e) Rand Value of municipal services completed - housing									
Score:	**V**	3.00	**R**	3.00	**T**	3.00	**P**	2.25	**I**	3.00

Data Limitations (Non-conformities):
Non-conformity 1: ***Precision***
MINOR - Potential errors may occur in that a standard subsidy amount is used for reporting purposes. Actual levels of subsidy may differ.
Strengths and Vulnerabilities:
Strength 1: ***Validity***
The data are based on well developed systems
Strength 2: ***Reliability***
Data collection has followed a consistent and easily auditable process
Strength 3: ***Timeliness***
Data are taken from a live data base
Vulnerability 1: ***Precision***
The Rand value is based on a standardized subsidy amount, not on actual expenditure.
Vulnerability 2: ***Source of Data***
The use of standardized subsidy amounts for reporting purposes could result in error
Recommendations:
Since the program is completed recommendations are not required

5. Data Quality Assessment Population Two

1.19. Summation of DQA Results for Population Two

Population two consisted of eight partners who report their data via a GMC held by Mega-Tech to USAID. In all cases the partners concerned were reporting GCC data and as the main purpose of the DQA was to assess SO6 data it was considered that on-site visits for these partners was not essential. The results of the DQA therefore review the data quality management systems of Mega-Tech by means of reviewing the data they receive from the reporting partners (Appendix M).

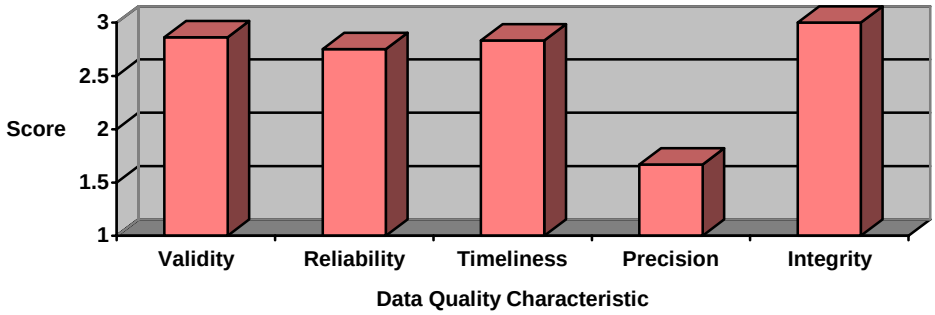

Figure 2: Average scores per quality characteristic for population two.

Risk to USAID/SA:

The risks presented by MT do not have their origins at MT – they are systems issues such as the management of indicators, especially the definitions of the GCC indicators. USAID/SA have also made many changes over the past few years. In terms of Mega-Tech data quality systems they demonstrate an overall low risk, with as is common, greatest risk presenting in the area of precision. The results suggest that the handling of partners through a GMAC reduces data quality risk.

The Agreement:

Mega-Tech is contracted to manage the partner agreements for USAID/SA. MT are responsible for collecting the data required from these contracted partners as well as the data quality of these partners.

Deliverables:

MT are responsible for collecting the data required from these contracted partners as well as the data quality of these partners. MT has to collate all the reports into one report for USAID/SA. The previous set of raw data was forwarded to USAID/SA, "as is" MT designed a standard report form for the partners contracted through MT – this form was introduced in October/November 2002. Partners attended a workshop in October 2002 where the form was explained to them.

On the latest agreements with partners, grantees were given a choice to decide which indicators to report against as per USAID/SA instructions. MT subsequently identified that the indicators reported on were not reflective of the grant programs and brought this to the attention of SO6. After a meeting to discuss the issue, SO6 asked MT to go back to the most recent SO6 (non-GCC) grantees and renegotiate more relevant indicators, which has been done. MT also designed a checklist – "Data Quality Questionnaire". This questionnaire has to be attached to the data form for the October 2003 data collection for the annual report to Washington. This process has just started so there are no records as yet on this new improved system.

Previous DQA's:

USAID/SA did a DQA on 12 December 2002 on "Data Quality Control: Global Climate Change (GCC) Related Data in Annual Report". The data used in the 2003 Annual Report was referenced.

Limitation(s) to Current DQA:

MT has made DQA visits to partners contracted through MT but no written records of these visits were made.

1.20. Analysis of Partners Reviewed

Eight partners were reviewed while auditing data quality at the MT offices – these were:

Agama Energy (AE)

This grant started on 26 April 2002 and ended on 30 November 2002. AE had to supply metered green electricity at WSSD and the grant value was R 1 031 814.

Cape Technikon (CT)

The project was based on a solar power system for electric sewing machines for a community without electricity. Three reports were handed in - 13 June 2002; 25 October 2002 and the final report on 25 February 2003. The expected result of the project was 5 community trainers; 60 seamstresses; 5 electric technicians and 10 trainee seamstresses.
At the time of audit the following was found:

- There was no evidence on-site that the data associated with the final report (25 February 2003) had been verified.
- The measurables are not specific in all cases.
- The final figures in the final report are not reflective percentages (measurables) of the expected results.

Food and Trees for Africa (FTFA)

This grant started on 23 April 2002 and was to end on 30 September 2003. The approved total was R 1 650 000 with a GMAC commitment of USD 150 000. The grant was amended on 30 May 2003 to end on 28 February 2004 with an eventual grant value of R 1 675 915 and a GMAC commitment of USD 200 000. FTFA has to report on two SO6 indicators semi-annually.
Four reports were handed in - dated 27 June 2002; period 31/03/2002 – 30/06/2002; period 01/07/2002 – 30/09/2002 and period 01/01/2003 – 31/05/2003. A "USAID Indicator Semi-Annual Report" for Apr – Sep as well as the FTFA official "Annual Review 2000-2001" was also seen.

International Institute for Energy Conservation (IIEC)

This grant started on 21 May 2002 and ended on 21 May 2003. IIEC had to assist with sustainable transport teams – capacity building, public awareness and the like. The grant value was R 454 080.

Lynedoch Development Foundation (LDF)

LDF had to do detailed technical plans, designs and legal approvals for a company to build a model eco-village. The grant started on 21/05/2003 and ended on 31/10/2002. The initial grant value was USD 58 000 but was amended to USD 65 751. A quarterly report dated 10 July 2002 was seen at audit.
At audit the following was found:

- There was no evidence on-site that the data associated with the closure of the project had been verified.
- The percent achievement of the milestones is not reported.
- There is no reference or report reflecting the expenditure against the budget.

Midrand Ecocity (ME)
This grant started on 30 April 2002 and ended on 30 November 2002. The initial grant value was USD 44 000 but was eventually amended to USD 50 616. ME had to research alternative ways of planning and organizing communities that value the consideration of sustainable development and also to provide a showcase for WSSD. Sections 7 and 8 on pages 4 and 5 (of 17) of Attachment 1 gives a method for verification of project results. ME reported against GCC 1.2 and GCC 3.1. A report dated 10/07/2002 was seen that demonstrated the percentage completion of the quantity report.

National Development Initiative for Social Welfare (NDISWE)
The grant was for a "Farm Scale Ethanol Production Plant – Demonstration Prototype". The grant started on 03/05/2002 and ended on 04/09/2002. The approved value was R 904 000 with a GMAC commitment of USD 87 000 (exchange rate was USD 1 = R 10.30). NDISWE also received R 1 200 000 from Department of Welfare

(national) and R 600 000 from the Department of Social Development (national). DEAT also donated R 6 000 for extra costs for WSSD. After an amendment was made to the grant, the total GMAC commitment increased to an eventual USD 91 134. NDISWE acted as a Project Coordinator and subcontracted out for technical aspects to Frameworks International CC and Aprocot. The grant was only for 4 months and yet they had to report on one SO6 indicator semi-annually.

At audit four reports were seen: "Farm Scale Ethanol Demo Plant – Grant Report" (dated 23 July 2002); "Operation and Production of the Village Scale Ethanol Production Plant"; USAID/GCC Grantee Data Collection Sheet (no date) and the "Grant Activity Completion Report" (dated 11 July 2003). All objectives were achieved – the plant is established in Silverton at the Agricultural Research Council.

Soweto Development Foundation (SDF)

This grant started on 03 May 2002 and ended on 03 April 2003. The grant value was R 1 100 640. SDF had to initiate urban greening and the development of conservation ethos in selected sites within Soweto. SDF had to report on one SO6 indicator in the format provided by GMAC.

1.20.1. Mega-Tech DQA 1 of 1

Result:	Handling of Data from Partners									
Score:	**V**	2.86	**R**	2.75	**T**	2.83	**P**	1.67	**I**	3.00

Data Limitations (Non-conformities):

Non-Conformity 1 Non Sampling Error:

MINOR – The data collection instrument - "USAID/GCC Grantee Data Collection Sheet" – is somewhat complicated and is open to interpretation by reporting partners.

Non-Conformity 2 Non Sampling Error:

MINOR – Definitions are not operationally precise.

Non-Conformity 3 Data Manipulation:

MINOR – MT does not check all data contained in the GDCS before forwarding to USAID/SA, missing data and also the accuracy of totals is a risk.

Non-Conformity 4 Consistency:

MINOR – The data collection system and/or instrument has been altered significantly during the reporting period. This introduces risk for collation of data at USAID/SA level.

Non-Conformity 5 Precision:

MINOR – The margin of error has not been determined by any of the partners that report through MT.

Non-Conformity 6 Precision:

MAJOR – It was not possible to determine the extent of error during this audit and thus it is not possible to determine whether such error can be reduced as a cost--effective and manageable interest.

Non-Conformity 7 Integrity:

MINOR – MT have not had an on-site visit from USAID/SA – no written record was seen at audit.

Non-Conformity 8 Data Source Type:

MAJOR– It is unknown if different data types are aggregated – this introduces an absolute risk.

Strengths and Vulnerabilities:

Strength 1: General

MT have tried to introduce a standardized data collection system for all the partners that report through them to USAID/SA – an idea that may be useful for all partners.

Recommendations:

Vulnerability 1 Transcription Error:

Although no transcription took place during the previous run of data collection from MT, there is a possible risk that USAID/SA should be made aware of. USAID/SA has requested that MT collate the data reported – there may be a potential for error with the next data collection when MT collates all the data from the various partners due to inherent data differences.

Vulnerability 2 Internal Quality Control:

The informal internal quality control systems at MT are not formalized resulting in low auditability.

6. Data Quality Assessment Population Three

1.21. Summation of DQA Results for Population Three

Two partners were to be assessed as part of population three. Of these no records could be located for one of the partners. The remaining partner would not undergo the DQA on-site due to the nature of the contractual relationship they have with USAID/SA. In terms of data quality the results of this DQA are shown below in Figure 3. The low score given to precision is based on the audit being unable to demonstrate the margin of error; likewise this applies for the integrity of the data.

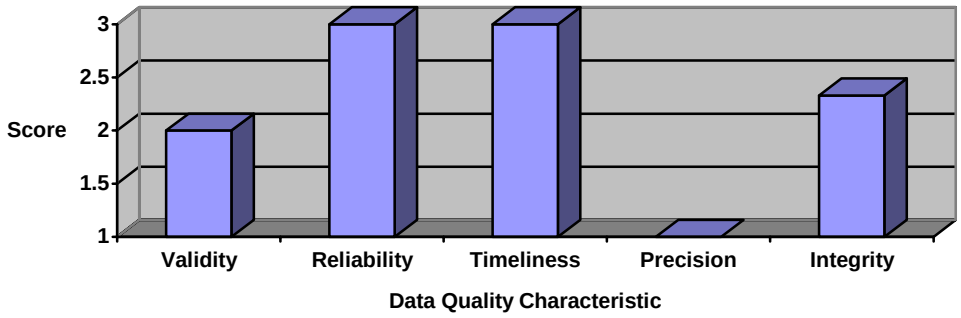

Figure 3: Average scores per quality characteristic for population three.

1.22. First Rand Bank

Risk to USAID/SA:

The greatest risk to USAI/SA in terms of this project lies in the issues of attribution to specific services and acceptability of the formulae constructs (Appendix N).

The Agreement and Deliverables

The First Rand Bank is a beneficiary of the Housing Guaranty Program. The funds are borrowed by the Bank against a US government guaranty, and used as loans to local governments to strengthen service provision. Under the conditions of the HG program, the funds must be used to benefit below median income groups.

The funds were not lent against specific projects, but go to the general revenues of the municipality, subject to a agreement that they will be applied to the capital costs of service provision. There is therefore no direct link between the funds lent and the number of households benefiting.

To overcome this difficulty, a formula was agreed by which the data could be generated. It would be assumed that the funds would be allocated to specific services in the proportions as budgeted for capital expenditure by the municipality for the year in which the loan was made. In order to derive numbers of households, the sums would then be divided by sums as agreed in a Project Implementation Letter of 15th December 2001, namely:

Electricity	R3 000 per household
Water	R1 500 per household
Sewage	R5000 per household
Roads	R4000 per household
Solid waste	varies

In addition, in order to derive a total number of persons benefiting, household size was set at four persons.

Reporting Validity:

Clearly attribution of funds leveraged in such a project is easy. The Bank borrowed $25 million and agreed to leverage South African funds equivalent to $140 million. Evidence of disbursements is provided in the form of local agreements with municipalities.

However, attribution of the funds to specific services and specific numbers of households cannot be verified. There was nothing in any of the agreements to prevent a municipality from using all the funds for a major water reservoir, for example. Thus while all citizens of the municipality might experience the continuation of a satisfactory service as a result of the expenditure, it is stretching a point to claim a certain number of households have experienced benefits.

Similarly, if all the money went to water, it is stretching a point to claim that households have experienced improved electricity services.

The logic of the formula is that if the funds had *not* been borrowed from FRB, then all the services would have received less income, and therefore electricity provision would have, indeed, suffered. However from the data point of view there is no doubt that the formula has the basic flaw of having no verifiable or auditable attribution in terms of the number of households served.

A second difficulty arises in terms of the formula that assumes that 50% of all the funds will benefit below median income families. This presents the same difficulties as the attribution to specific services.

The matter is more easily audited when reporting against funds leveraged, but the question of attribution to specific services remains unverifiable.

Within the limits of the formula, the observation is that the numbers used to derive households served are quite conservative when considering costs for below median income households. Thus, if the funds were exclusively used for servicing land for RDP housing, the data generated would certainly be an under-estimate of the number of households. On the other hand, if upper income households were the beneficiaries, the errors would be compounded, as the costs would be too low and the attribution would be more seriously skewed.

1.22.1. FRB DQA 1 of 2

Result:	IR 6.2.2: Number of households assisted to obtain shelter/urban services through the provision of credit and subsidies to low income communities									
Indicator:	Number of HDP households assisted to obtain new or improved shelter/urban services through the provision of credit and subsidies									
Score:	**V**	2.00	**R**	3.00	**T**	3.00	**P**	1.00	**I**	2.33
Data Limitations (Non-conformities):										

Non-conformity 1: Validity
MAJOR - The formula used has two major weaknesses – while funds are being provided for general municipal use they are attributed to specific households and specific services. There is, in practice, no demonstrable link between the two.
Non-conformity 2: Validity
MINOR - There is no measurement of the degree to which the services provided are environmentally sustainable.
Non-conformity 3: Validity
MINOR - Errors in using local government data are possible.
Non-conformity 4: Validity
MAJOR - Significant risk is introduced with manipulations of secondary and tertiary data sources.
Non-conformity 5: Validity
MAJOR - The data are based on formulae and cost assumptions regarding the expenditure for specific services. Such cost data has not been verified in the field. It may be unduly conservative, but the fact has not been verified
Non-conformity 6: Precision
MAJOR - There are insufficient direct linked between the data and the results claimed to establish the margin of error
Strengths and Vulnerabilities:
Strength 1: Reliability
The data are consistently collected and applied in a transparent manner
Strengths 2: Timeliness
The data are available on an immediate basis.
Vulnerability 1: Data Source Type
The derivation of the population/household data has not been checked.
Recommendations:
Recommendation 1: Validity
The data should not be disaggregated into specific services unless the funds were specifically allocated for such a service.

1.22.2. FRB DQA 2 of 2

Result:	IR 6.2.1 Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision								
Indicator:	6.2.1 (a) – (d) Indicator description: Total Rands in millions provided, including funds for new or improved housing or services leveraged for HDP households								
Score:	**V**	1.92	**R**	2.85	**T**	3.00	**P**	1.00	**I**	2.33
Data Limitations (Non-conformities):										

Non-conformity 1: Validity

MAJOR - There is no demonstrable link between the funds spent and the results claimed in terms of HDP households assisted to obtain shelter/urban services.

Non-conformity 2: Validity

MINOR - Errors in using local government data are possible.

Non-conformity 3: Validity

MAJOR - Significant risk is introduced with manipulations of secondary and tertiary data sources.

Non-conformity 4: Reliability

MINOR - The data reported cannot be objectively verified

Non-conformity 5: Precision

MAJOR - The margin of error cannot be demonstrated to be less than the expected change being measured as the nature of the interpretation makes any verification impossible

Strengths and Vulnerabilities:
Strength 1: ***Timeliness***
The data are available on an immediate basis
Vulnerability 1: ***Precision***
There is no direct link between the expenditure claimed and the percentage of expenditure going to HDP households.
Recommendations:
Nil

1.23. Kutlwanong Civic Integrated Housing Trust

Risk to USAID/SA

No DQA could be performed on the Kutlwanong Civic Integrated Housing Trust as it was not possible to trace the Trust to any of the physical addresses or contact information supplied at the beginning of the audit. It was also not possible to review any documentation with regards this partner as no documentation other than the original agreement was traceable, at the USAID/SA offices in Pretoria. The risk to USAID/SA associated with data from this partner must therefore be considered absolute.

7. Performance Indicator Quality Assessment

The purpose of this exercise was to review quality issues pertaining to the indicators for Strategic Objective 6: Increased access to shelter and environmentally sound services.

The performance indicators are assessed below against four indicator quality characteristics - directness, objectivity, practicality, and adequacy. In each case, the results of the qualitative assessment are given together with a discussion of data quality issues. In addition, recommendations related to each specific indicator are made, where relevant, and which cover such areas as: proposed indicator definition, rationale, frequency of data collection, methodology, responsibility for data collection, target, and data limitations.

1.24. Indicators 6.1: (a) – (e) Rand Value of Municipal Services

Performance Indicator: 1 (a) – (e) Rand value of municipal services (SO Level Indicator)

Indicator Description: SO level indicator which tracks the Rand Value of new or improved municipal services (water sewage, solid waste, electrical and gas connections, housing) stated leveraged through USAID support programs. This is an output indicator that indirectly but reliably measures access to the services stated.

Directness:	**Uneven**	***Objectivity:***	**Medium**	***Practicality:***	**Medium**	***Adequacy:***	**Medium**
DIRECTNESS							
Uneven: The provision of funding has a positive correlation to increased access to shelter and urban services, but funding does not have any correlation with the degree to which services are environmentally sound.							
OBJECTIVITY							
Medium: The lack of a clear definition about the word leveraged, and lack of guidelines regarding the parameters for claiming leveraging undermine the objectivity of the data collected against this indicator.							
PRACTICALITY							
Medium: Since the indicator requires funds leveraged to be identified, the data must necessarily be collected from third parties. This introduces risks in terms of the reliability of the data. In other respects the indicator is a practical and easily administered one							
ADEQUACY							
Medium: By measuring expenditure at specific moments in time, the data are cross sectional, not longitudinal. This limits the indicator's use as a measure of progress.							
DATA QUALITY ISSUES							
Primary data collected by partners is of a high quality, but difficulties arise in some cases in that the funds leveraged are calculated by the SO team, not the partners. This introduces the possibility for minor error. The larger issue is that the indicator measures funds leveraged, but there is no definition of how leveraged can be defined: this raises the difficulty of assessing the true impact of USAID assistance.							
RECOMMENDATIONS							
Proposed Indicator: Rand value of funds leveraged for shelter and urban services.							

Definition: Indicator tracks the Rand Value of new or improved municipal services stated leveraged through USAID support programs. Funds leveraged must be verifiable as a matter of fact and degree as a direct outcome of USAID financial support so that a direct financial ratio can be established. E.g. technical assistance to a program cannot be construed as leveraging funds.
Rationale: Two main modifications are proposed. 1. Definition of the word leveraged to limit it to financial leveraging 2. Omission of reference to environmentally sound services, which should be subject to different reporting indicators
Frequency of Data Collection (New and current partners): 1. Baseline data to be collected at the inception of each activity with any new partner. 2. Thereafter reporting to be at biannual intervals (end March and end September).
Methodology: The existing methodology is satisfactory, except that the system and the input data from partners by which data is aggregated in the SO office must be fully documented.
Responsibility for Data Collection: Active partners for input data.
Target: Increase in the value of funds provided for the provision of shelter and urban services.
Data Limitations: Some partners do not report on the value of shelter units provided. This is not satisfactory, and can be remedied without significant burden on the partners concerned

1.25. Indicators 6.2 (a) – (e): Number of Households Receiving Municipal Services

Performance Indicator: 2 (a) – (e) Number of households receiving municipal services (SO level indicator)

Indicator description: Number of households with new or improved service*. Indicator tracks the number of households receiving municipal services provided by USAID supported programs

***In the PMP each indicator sheet specifies a separate service, namely water, sewage, solid waste, electrical and gas connections, housing.*

Directness:	**Uneven**	***Objectivity:***	**Poor**	***Practicality:***	**Medium**	***Adequacy:***	**Uneven**
DIRECTNESS							
Uneven: The indicator is direct in terms of access, but not in terms of environmental soundness. It also lacks precision in that the indicator covers both new and improved urban services. A lack of definition of improved services means that a de minimis intervention in an urban service can be claimed to bring improved services to many households							
OBJECTIVITY							
Poor: There is confusion regarding the concepts of new and improved services in the indicator, thus measuring more than a single phenomenon. Also the indicator requires that households benefiting from services provided by "USAID supported programs" should be included. The words "supported by" require definition so as to obtain a verifiable relationship between USAID support and the provision of a service.							
PRACTICALITY							

Medium: For some partners the practicality is high in that the number of units is the basis for their own data management. In the field of urban services, the data must be manipulated and questions of definitions and attribution must be addressed.
ADEQUACY
Uneven: The data are adequate in terms of the measurement of housing units completed. They are not adequate in terms of improved services.
DATA QUALITY ISSUES
There are three difficulties in terms of data issues. 1. The definition of the degree to which the data used can be attributed to "USAID supported programs" 2. The attribution of expenditure of services to specific households 3. Use of the term household. All data collected would appear to be on the basis of a dwelling, not household. Since it cannot be assumed that every dwelling has a single household this leads to under-reporting. Furthermore, actual household size is never measured, and use of standardized census-derived values may be misleading.
RECOMMENDATIONS
Proposed Indicator: Number of shelter units constructed through programs supported by USAID.
Definition: Indicator tracks the number of units developed through USAID supported programs. Units reported must be verifiable as a matter of fact and degree as an outcome of USAID financial support.
Rationale: Two main modifications are proposed. 1. Definition of the word "provided" to ensure that clear attribution is possible 2. Omission of reference to environmentally sound services, which should be subject to different reporting indicators
Frequency of Data Collection (New and current partners): 1. Baseline data to be collected at the inception of each activity with any new partner. 2. Thereafter reporting to be at biannual intervals (end March and end September).
Methodology: The existing methodology is satisfactory, except that the system and the input data from partners by which data is aggregated in the SO office must be fully documented.
Responsibility for Data Collection: Active partners for input data.
Target: Increase in the value of funds provided for the provision of shelter and urban services

1.26. Intermediate Result: IR 6.1.1

Intermediate Result: IR 6.1.1: Improved policy environment for facilitating access to shelter and urban service

Performance indicator: Impact of policy frameworks on implementation constraints upon housing and municipal services development

Indicator description: The indicator is a number derived from a matrix, based on the degree to which policy constraints on housing municipal services delivery are being removed. A team of specialists undertakes compilation of the results.

Directness:	Poor	*Objectivity:*	Poor	*Practicality:*	Poor	*Adequacy:*	Poor
DIRECTNESS							
Poor: The construction of a matrix to overcome the difficulties of measurement of non-quantifiable variables is an imperfect tool. There is no direct relationship between the phenomena measured and the results reported.							
OBJECTIVITY							
Poor: It cannot be objectively verified as a measurement of the result. Inherently subjective judgments have to be used to construct the matrix.							
PRACTICALITY							
Poor. The data used to construct the matrix require special expertise. There is no certainty that those operating the matrix from reporting period to reporting period (two years) will use the same criteria.							
ADEQUACY							
Poor. The matrix does not distinguish between policy changes induced by USAID supported programs and those derived from other programs.							
DATA QUALITY ISSUES							
The matrix is basically an un-auditable device due to its reliance on informed judgments by experts.							
RECOMMENDATIONS							
It is proposed that narrative reporting should be used for this subject.							

1.27. Intermediate Result: IR 6.1.2

Intermediate Result: IR 6.1.2: Improved policy environment for facilitating access to shelter and urban services

Performance indicator: Impact of policy frameworks on implementation

Indicator description: Indicator tracks totals of subsidies issued by Department of Housing, Number of areas covered by MIIF, Number of builders registered with the National Home Builders Warranty scheme, Rand value of credit made available to retail banks by NHFC and number of subsidies facilitated by the PHPT.

Directness:	Poor	*Objectivity:*	Poor	*Practicality:*	Poor	*Adequacy:*	Poor
DIRECTNESS							
Poor. The indicator measures outputs which cannot be ascribed to changes in the policy environment							
OBJECTIVITY							
Poor. Interpretation is made difficult by the fact that multiple sectors are included in a single indicator, and the relationship between USAID programs and the outputs measured is not a direct one.							
PRACTICALITY							
Poor. The data has to be collected from contractual partners, and the two yearly interval for data collection is too infrequent to inform management decisions							
ADEQUACY							
Poor. The indicator is not a measure of progress of programs directly attributable to USAID, nor of policy change due to low correlation. It is not possible to measure attribution from the data.							

DATA QUALITY ISSUES
The indicator units many different units of measure, e.g. number of subsidies issued, Rand value of credit made available to banks by the NHFC, number of builders registered etc. It is thus aggregating data without defining benchmarks for progress.
RECOMMENDATIONS
Proposed Indicator: The indicator should be abolished.

1.28. Intermediate Result: IR6.1.3

Performance indicator: IR6.1.3 Shelter units completed

Indicator description: Approved subsidies or credits for shelter units through SO6 supported programs

Directness:	**Medium**	*Objectivity:*	**Medium**	*Practicality:*	**High**	*Adequacy:*	**Medium**
DIRECTNESS							
Medium. The indicator is direct in terms of the number of units completed, but shelter units completed is not the direct result of approval of subsidies or credits.							
OBJECTIVITY							
Medium. Measurement of the subsidies or credits approved is not directly linked to shelter units completed. There is a risk of double counting as units may receive subsidies and credits.							
PRACTICALITY							
High. The practicality is high for most partners in that the number of subsidies/credits is the basis for their own data management.							
ADEQUACY							
Medium. The data are adequate in terms of the measurement of the number of subsidies/credits but this does not measure units completed							
DATA QUALITY ISSUES							
The definition of the degree to which the data used can be attributed to "USAID supported programs" Risk of double counting, as partners report separately, and the SO6 office aggregates the data. It is quite likely that more than one partner will be reporting on a single scheme.							
RECOMMENDATIONS							
Proposed Indicator: This indicator is so close to 2. (e) that consideration should be given to abolishing it, with suitable revisions to 2 (e) as proposed above.							

1.29. Intermediate Result: IR 6.2

Intermediate result: IR 6.2 Previously ineligible households developers builders and municipal service providers obtaining access to credit

Performance indicator: 1 - Rand Value of credit and subsidies obtained for households for HDP shelter and urban services provision

Indicator description: "Total Rands in millions provided, including funds for new or improved housing or services leveraged for HDP households"

Directness:	**Poor**	*Objectivity:*	**Medium**	*Practicality:*	**Medium**	*Adequacy:*	**Medium**
DIRECTNESS							
Poor: the provision of funding has a positive correlation to increased access to shelter and urban services, but the data collected do not reflect access to credit by the parties specified in the result being reported against							
OBJECTIVITY							
Medium. The lack of a clear definition about the word leveraged, and lack of guidelines regarding the parameters for claiming leveraging undermine the objectivity of the data							
PRACTICALITY							
Medium. Since the indicator requires funds leveraged to be identified, the data must necessarily be collected from third parties. This introduces risks in terms of the reliability of the data. In other respects the indicator is a practical and easily administered one							
ADEQUACY							
Medium. By measuring expenditure at specific moments in time, the data are cross sectional, not longitudinal. This limits the indicator's use as a measure of progress.							
DATA QUALITY ISSUES							
Primary data collected by partners is of a high quality, but difficulties arise in some cases in that the funds leveraged are calculated by the SO team, not the partners. This introduces the possibility for minor error. The larger issue is that the indicator measures funds leveraged, but there is no definition of how leveraged can be defined: this raises the difficulty of assessing the true impact of USAID assistance.							
RECOMMENDATIONS							
Proposed Indicator: As currently operated this indicator is identical to SO Indicators 1(a) – (e) and is therefore redundant							

1.30. Intermediate Result: IR 6.3

IR 6.3: Increased non-credit forms of assistance made available to the historically disadvantaged population

Intermediate Result: IR 6.3.1: Local Authorities with improved financial management & service delivery capacity

Performance indicator: Local Authorities with improved financial management capacity and service delivery capacity.

Indicator description: The indicator tracks the number of stages that local authorities have advanced to address financial management, water/sanitation/solid waste, transportation, HIV/AIDS or urban renewal, as per chart in unit of measure

Directness:	**Poor**	*Objectivity:*	**Medium**	*Practicality:*	**High**	*Adequacy:*	**Poor**
DIRECTNESS							
Poor: There are no objectively verifiable links between the events reported							
OBJECTIVITY							

Medium. The objectives are clear, as defined by the performance indicator. However, the indicator description lacks objectivity in that graduation from one stage to the next does not necessarily indicate improved management capacity. Furthermore, the data reflecting the results do not measure increased capacity, but only participation in the program.
PRACTICALITY
Good. The primary data in the indicator are collected very simply
ADEQUACY
Poor. The data represent only Local Authorities participating in the program, not a measure of their increased capacity.
DATA QUALITY ISSUES
It is not clear how the indicator is expected to operate in practice due to the design of the form in the PMP. If partners are requested to report on the stages reached, the risk is that they will not reflect any increase in capacity, but simply an administrative transition. Even if the transition from one designated stage to the next were captured objectively, it is not clear how the fact is recorded in the indicator itself.
RECOMMENDATIONS
Proposed Indicator: Local Authorities assisted to develop improved financial management capacity and service delivery capacity.
Definition: A measure of the degree to which Local Authorities have received on-the-job support and capacity building in financial management and service delivery. One unit represents completion of one project cycle in one of five sectors. Each component of the cycle is recorded as 0.25 points – thus completion of a complete cycle within a Local Authority scores one point. The maximum score per Local Authority would therefore be 5.
Rationale: The indicator can be used as it is currently designed, but the form must be re-designed so that the data captured reflects the transition made by each Local Authority from one stage to the next.
Frequency of Data Collection (New and current partners): As generated.
Methodology: The current method of capacity building is to appoint consultants to perform each stage of the work. Consultants would therefore be required to meet output criteria: on official adoption of their work the task could be declared complete. This would be the source of primary data. Consultants would be appointed as capacity builders not to undertake the work directly.
Responsibility for Data Collection: Consultant agreements
Target: Capacity building of municipalities in financial management and service delivery
Data Limitations: Arriving at appropriate benchmarks for graduation from one stage to the next may be difficult. However, official approval of projects by funding agencies would probably be an adequate proxy.

8.8 Intermediate Result: IR 6.4

IR 6.4: Improved capacity to apply sustainable participatory environmental principles to local level urban development

Performance indicator: Emissions of carbon dioxide equivalents avoided

Indicator description: The indicator captures the emissions avoided disaggregated into three categories:

1. Carbon dioxide emissions avoided through renewable energy activities
2. Carbon dioxide emissions avoided through end use energy efficiency
3. Carbon dioxide emissions avoided through energy efficiency improvements in generation, transmission and distribution (including new production capacity)

Direct-ness:	**Poor**	*Object-ivity:*	**Medium**	*Practic-ality:*	**High**	*Ade-quacy:*	**Medium**

DIRECTNESS

Poor: There are no commonalities between the indicator and the results measured. The former refers to "improved capacity to apply . . .principles", the measurement refers to reductions in carbon dioxide emissions. The indicator would seem to be relying on measures which would generate increased capacity, such as training and technical assistance. The later is related to the adoption of technologies which reduce reliance on fossil fuels.

OBJECTIVITY

Medium There are weaknesses in the objectively verifiable nature of the data due to the necessity for formulae to be used to calculate the data, and the lack of empirical data to support the formulae.

PRACTICALITY

Good. The primary data in the indicator are collected very simply

ADEQUACY

Medium: The indicator is a satisfactory definition of progress, but attribution is not defined

DATA QUALITY ISSUES

There are two issues. First the question of the use of formulae to relate specific solutions, e.g. converting energy saving solutions to reductions in carbon emissions, which have not been verified by supporting research (e.g. does the introduction of improved insulation reduce fuel consumption or result in a higher standard of heating?). Second attribution is not defined, so it is not clear whether all the reductions are produced by expenditure of USAID funds, or funds leveraged from them.

RECOMMENDATIONS

Proposed Indicator:
Carbon dioxide emissions avoided

Definition: Carbon dioxide emissions avoided as a result of USAID funded programs, or funds leveraged through USAID funded programs, in the following categories

1. Through renewable energy activities
2. Through end use energy efficiency
3. Through energy efficiency improvements in generation, transmission and distribution (including new production capacity)

Rationale:
The indicator can be used as it is currently designed, but attribution must be made more clearly.

Frequency of Data Collection (New and current partners):
As generated.

Methodology: The existing methodology, which is also used for GCC reporting, is satisfactory.

Responsibility for Data Collection:
Consultant agreements

Data Limitations:

1. Attribution should be clarified.
2. Research to determine the accuracy of the assumptions of the formulae should be considered.

8. The Strategic Objective and Performance Indicator Relationships

1.31. Organizational Framework (OF) Description

In undertaking the assessment of both indicator quality data quality, the Team considered the development model embodied in the strategic framework, and the relationships between the various levels of interactions, such as impacts, outcomes, outputs, activities and inputs. Knowledge of the model and these relationships helped to clarify the characteristics of the indicators, the nature of the data being collected by partners, and partner reporting responsibilities. An Organizational Framework (OF) is being used to present this information. The OF is simply another way of representing the SO6 results framework, albeit one that is "stretched" to include the full results "chain levels" embodied in the SO (inputs, activities, outputs, outcomes and impacts). The respective chain levels correspond to the results framework, as follows:

- Impact level corresponds to the SO
- Outcome level corresponds to IRs
- Output level corresponds to partner performance indicators
- Activity level corresponds to operational activities of implementing partners
- Inputs correspond to important partner data collection items or "data collection points."

In addition, the OF is divided into two broad categories: the external environment (influenced by factors outside partner activities) and internal environment (the local environment in which program partners operate). Partner data collection and reporting responsibilities are represented by blue ovals.

For comparative purposes, two OFs are presented. Figure 4 reflects the recommendations of the Team.

1.32. Current Indicator Relationships - OF

The current construction of the SO PMP has allowed for the majority of the measurable indicators to lie at the impact level of the SO (10 of 17 indicators). As all 10 of these indicators are constructed as output indicators, which result from the activity of a partner, this relationship is inappropriate. By and large the Rand value based indicators can be objectively verified, however the disaggregation into the various municipal services, as required by Washington, is inappropriate for the South African environment. 'Household' based data relative to services is inevitability speculative as opposed to 'shelter' data, which can be measured quantitatively.

At the Outcome level there are four intermediate results of which one measures impact (IR6.1). This relationship is also inappropriate as the nature of data collection, collation and analysis for the indictors for this result do not support its being placed at the outcome level. As there are manipulated calculations based on the results and outcomes of other activities this result and its indicators must be a measure at the impact level, which remains separate to partner activities.

The indicators for two of the Intermediate Results (IR6.2 &IR6.3) are an operational repeat of indicators at the SO level due to the discrepancies that exist in the definitional issues. This means that in practice they are interpreted and measured as the same entity. This adds no value to the ultimate measurement of the achievement of the SO.

The last Intermediate Result (IR6.4) is the only one, which is aimed at measuring the environmental component of the SO. It construction and use of terminology makes the indicator associated with the result non-measurable. Furthermore the description given to the indicator is unrelated to the IR.

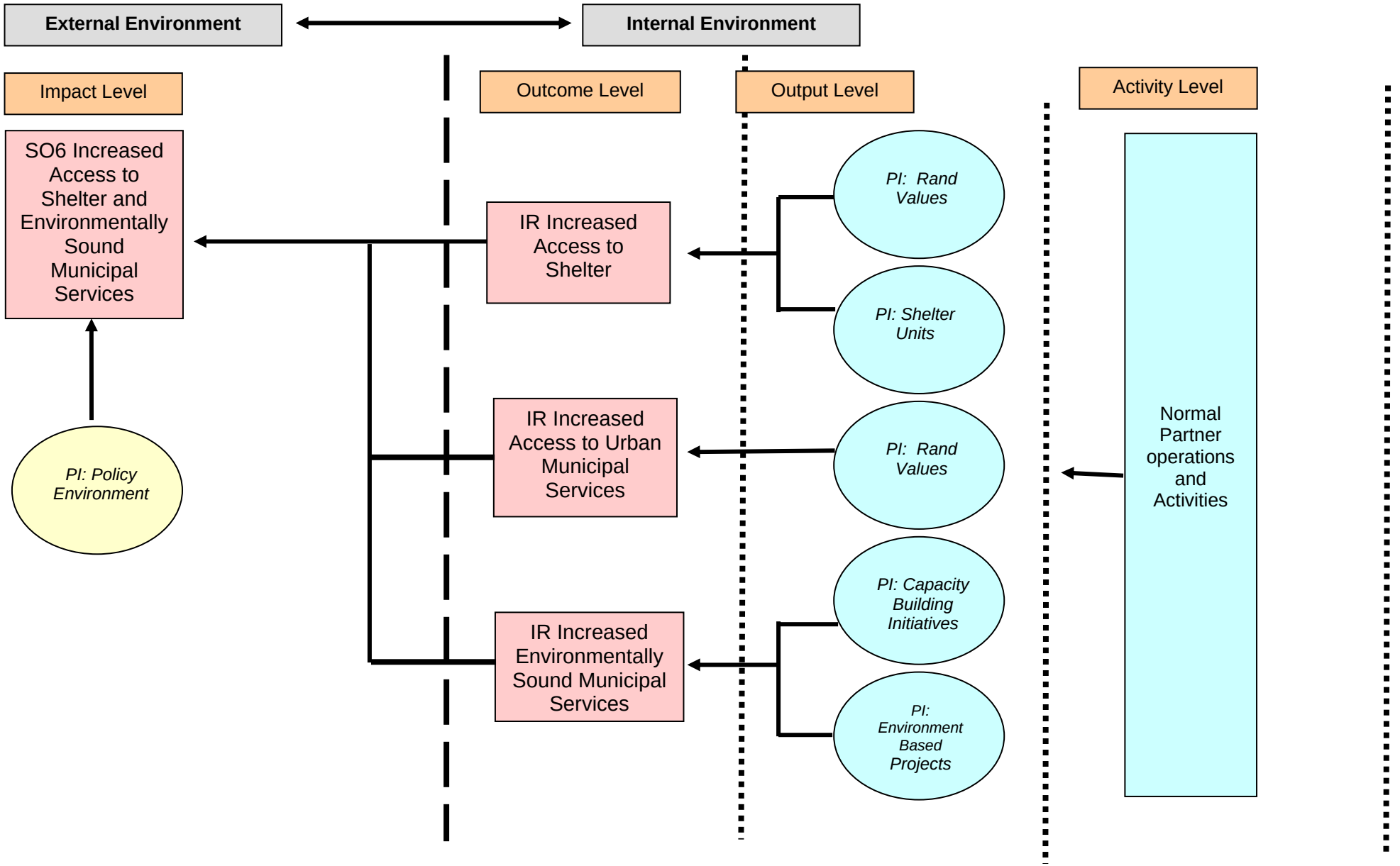

Figure 4. Proposed Organizational Framework for Strategic Objective Six

1.33. Proposed Indicator Relationships - OF

The first task that must be achieved is that the operational relationships between the various levels within the OF of the PMP must be re-assessed for validity. The number if indicators should be reduced so that unnecessary duplications and overlaps are eliminated. The two separate elements of the SO; being "Increased Access to Shelter" and "Environmentally Sound Municipal Services" must be addressed in their own rights for measurement validity. The reason for this is the inherent differences that will exist between the quantitative nature of the first component of the SO and the qualitative nature of the second half of the SO.

No more than two to three indicators per SO element should be included at the IR (outcome level). These need to be direct measures which can be shown to have a high correlation with the impact level, e.g. Increased financial assistance made available to municipalities for urban services. Such an indicator can be managed at the output level by partners as a specific Rand Value measure of funding granted etc.

1.34. Systems Level Issues

The greatest risk presented to USAID/SA in terms of the management of this SO PMP, as observed during this DQA, is related to the systems integrity and auditability. In practice this means that vulnerabilities have been demonstrated between the strategic planning, operational management and monitoring of activities related to the PMP. The lack of traceable and accessible records, at all levels, leaves the data reported against this SO open to question. Inconsistencies of the internal management of aggregation of data were demonstrated during audit. It was not possible to backtrack any decision-making or audit trail for these inconsistencies.

In essence this DQA has demonstrated that at the systems level there is some degree of failure to implement a Plan, Do, Check and Act management cycle (Deeming Cycle) which, within this type of environment is usually considered best practice.

9. Summary of Recommendations

1.35. Data Quality Assessment Level

1.35.1. Partner Level

a. Each partner should have a data quality plan which outlines how the partner manages data to ensure that the least amount of risk is introduced into the system for each of the data quality attributes.

b. Partners should identify the inherent strengths and vulnerabilities of their data management system and report these to USAID/SA as part of their routine reports.

c. Partners must ensure that they maintain consistency of data management and any changes in data handling methodologies must be reported to USAID/SA.

d. All data reported by partners must be backed up by a verifiable audit trail.

1.35.2. USAID Level

a. Data quality should be an inherent component of the acceptance criteria for the contracted deliverables of reporting partners.

b. USAID/SA should consider additional capacity building initiatives, similar to those implemented by Mega-Tech to ensure that partners are able to manage data quality issues.

c. Partners should only be contracted to report quantitative data, which can be clearly demonstrated to have a high Pearsons Correlation Coefficient for validity against the indicator.

d. When contracting partners, the data reporting frequency must take cognizance of allowing for the early identification and management of any inherent data quality risks.

e. Where partners have indicated a change in data handling methodology USAID should ensure that they have documented, and calculated, the effect of the changes on the quantitative result.

f. Margins of Acceptable Error should be set at USAID/SA level for quantitative indicators, as a whole, and the incoming data assessed for the potential of being an outlier.

g. Where data from a partner has an established risk this should be accompanied by an USAID/SA risk management plan.

h. Outsource regular data quality audit of partners on the basis of a GMAC.

1.36. Performance Indicator Quality Assessment Recommendations

a. Reconstruct PMP to reflect appropriate relationships between indicators and their position within the OF.

b. Reduce number of indicators.

c. Performance Manage the PMP on the basis of a Deeming Cycle.

d. Construct and implement a systems level Risk Management Plan for data already reported against the SO6 PMP and for which systems level vulnerabilities have been demonstrated.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX A: GUIDELINES FOR QUALITY AND/OR ENVIRONMENTAL MANAGEMENT SYSTEMS AUDITING

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

This report was prepared under Mega-Tech, Inc.'s prime contract with USAID and addresses USAID/South Africa's Strategic Objective No. 6: Increased access to shelter and environmentally sound municipal services

ISO 19011:2002 - A COMBINED AUDITING STANDARD FOR QUALITY AND ENVIRONMENTAL MANAGEMENT SYSTEMS

Gary L. Johnson
Environmental Engineer
U.S. Environmental Protection Agency
Quality Staff (MD-34)
Research Triangle Park, NC 27711

Abstract

In a precedent-setting decision in 1998, the International Organization for Standardization (ISO) directed ISO Technical Committee (TC) 176 on Quality Management and ISO TC 207 on Environmental Management to develop jointly a single guideline standard for auditing quality and environmental management systems. When approved, this standard would replace ISO 10011-1, ISO 10011-2, and ISO 10011-3 on quality auditing and ISO 14010, ISO 14011, and ISO 14012 on environmental auditing. A Joint Working Group (JWG) was established comprising experts from both TC 176 and TC 207 to develop the new standard, ISO 19011, Guidelines on Quality and/or Environmental Management Systems Auditing, and to incorporate lessons learned from efforts to improve compatibility between ISO 9001/9004 and ISO 14001/14004, the standards for quality and environmental management systems, respectively. Work is proceeding on the development of ISO 19011 with an expected completion in the summer or fall of 2002.

INTRODUCTION:

This paper discusses ISO 19011:2002, *Guidelines on Quality and/or Environmental Management Systems Auditing*, an international consensus standard currently under development that provides guidance on auditing quality management system as well as environmental management systems. The paper includes a description of the standard, a discussion of relevant issues addressed during its development, and a summary of its current status. Following the approval of the ISO 14001 and ISO 14004 environmental management systems (EMS) standards and the start of a revision to the ISO 9000 quality management systems (QMS) standards, there was

considerable interest by ISO in increasing the compatibility between the EMS and QMS standards. Early in the discussions, it became clear that the similarities among the existing EMS and QMS auditing standards would make them a prime candidate for integration into a single standard.

A Joint Working Group (JWG), composed of participants from ISO/TC 176 on Quality Management and ISO/TC 207 on Environmental Management was created by ISO to develop the new standard. The JWG would have co-conveners, one from TC 176 and one from TC 207, and experts would be drawn from both technical committees. Because this venture had never been attempted by ISO before, the ground rules for operating the standard-setting process also had to be revised. Both TC 176 and TC 207 would participate fully in the process. Ballots would be sent to national member bodies for both technical committees, but ISO's rule of "one country, one vote" would require that both TC's agree on the vote for a particular ballot. Otherwise, a country's vote would not be counted. To ensure that a consensus position is reached in the USA, the U.S. Technical Advisory Groups (TAGs) to TC 176 and TC 207 formed a Liaison Group with representatives from TAG 176/Subcommittee 3 on Quality Auditing and TAG 207/ Subcommittee 2 on Environmental Auditing to formulate the U.S. position on ballots.

In November 1998, the first meeting of the JWG to develop a common auditing standard was held in The Hague, The Netherlands. Experts from TC 176 and TC 207 representing 34 countries attended that meeting with the purpose of charting the development process for the new standard. From the outset, the stronger experience was with the quality auditing standards. The environmental auditing standards had been published only for a little over two years and there wasn't much experience in their use. While very similar, there were some distinct differences between the quality auditing philosophy and that of environmental auditing. Issues getting early attention included auditor competency, usability by small-to-medium enterprises (SMEs) and developing countries, and the structure of the standard.

By the spring of 1999 and the second JWG meeting in Buenos Aires, Argentina, an initial Working Draft (WD.1) of the standard had emerged. Discussions were held at the TC 207 meeting in Seoul in June 1999 and at the TC 176 meeting in San Francisco in September 1999, which resulted in the first Committee Draft (CD.1) of ISO 19011. CD.1 was balloted in late 1999 and more than 1400 comments from 35 countries were received by the JWG Secretariat by the end of February 2000.

The JWG met in Berlin, Germany, in March 2000, to address the comments on CD.1. The JWG was divided into two sub-groups, one to address comments on the structure and process aspects of the standard, and one to address the comments on auditor competency. Each sub-group had about half of the comments. After considerable debate, the draft for CD.2 emerged and was balloted for comments in April 2000. The comments were received in August and were addressed by the JWG in Cancun, Mexico in September. The Cancun meeting produced CD.3 which subsequently distributed for comments in late fall 2001. The extensive international

comments on CD.3 were addressed in Sydney, Australia, in March 2001, and general consensus was reached on enabling ISO to issue a Draft International Standard (DIS) on ISO 19011 for a five month ballot among the ISO member countries. While the DIS stage generally means that most issues have been resolved, there are still some concerns about parts of ISO/DIS 19011 that may yield extensive international comments. The JWG is tentatively scheduled to meet in Vancouver, Canada, in early 2002 to review any comments received on the DIS. When the DIS is approved the standard will be elevate to the Final Draft International Standard (FDIS) stage and issued for a straight "yes or mo" ballot. Approval of the FDIS will result in ISO issuing the finalized standard. The goal is to publish ISO 19011 as an international consensus standard by the fall of 2002.

PURPOSE OF THE STANDARD:

ISO 19011 is intended to provide guidelines for auditing ISO 9001-based quality management systems (QMS) and ISO 14001-based environmental management systems (EMS); however, it will also be sufficiently general such that it can be applied to any QMS or EMS. The standard will replace the following current ISO standards:

- ISO 10011-1, -2, -3, *Guidelines for Auditing Quality Systems*
- ISO 14010, *Guidelines for Environmental Auditing - General Principles*
- ISO 14011, *Guidelines for Environmental Auditing - Audit Procedures - Auditing of Environmental Management Systems*
- ISO 14012, *Guidelines for Environmental Auditing - Qualification Criteria for Environmental Auditors*

ISO 19011 reflects the changes made to ISO 9001:2000, *Quality Management Systems - Requirements*, which was issued in December 2001, including the new business model for the standard. ISO 19011 is intended to apply to both internal and external auditing, and may be used as part of auditor certification and training.

STRUCTURE OF ISO 19011:

The structure of ISO 19011 is as follows:

0. Introduction
1. Scope
2. Normative References
3. Terms and Definitions
4. Principles of Auditing
5. Managing an Audit Program
6. Audit Activities
7. Competence of Auditors

The standard includes several diagrams and help boxes to aid users in understanding and using the guidance.

ISO 19011 is a guideline standard which means its use is not mandatory unless it is invoked as part of a multiple party agreement, such as contract or other legal agreement. As a guideline standard, its implementation is generally not auditable because the elements of the standard are not requirements and because there may be others ways of accomplishing the same objectives.

ISO 19011 is generally organized as follows: Clause 0, Introduction, assists the reader in understanding the reason for the standard and who might use it. Clause 1, Scope, defines the scope and applicability of the standard which extends beyond QMS and EMS auditing. Clause 4 provides general some principles on auditing to aid first-time users. Clause 5 provides guidance on establishing, managing, and evaluating different types of audit programs. Clause 6 addresses the process of planning, conducting, and evaluating individual audits within a specific audit program. Clause 7 addresses issues pertaining to auditor competence, including their initial selection and on-going evaluation.

THE AUDIT PROCESS:

Clause 4 - Principles of Auditing:

The standard provides a brief summary of auditing principles in Clause 4. These principles should be used to drive the establishment and implementation of the audit process for an organization. Key among the principles cited for auditor behavior are:

- **ethical conduct** -- the foundation of professionalism,
- **fair presentation** -- the obligation to report truthfully and accurately, and
- **due professional care** -- application of reasonable care in auditing.

Two other principles of auditing relate to the audit process primarily. They are

- **independence** -- the basis for impartiality and objectivity of the audit conclusion, and
- **evidence** -- the rational basis for reaching audit conclusions.

Clause 5 - Managing an Audit Program:

Clause 5 provides guidance for those who need to establish and maintain an ongoing set of audits for an organization. The standard utilizes the Plan-Do-Check-Act cycle to define the audit program. Some of the key actions addressed are:

- establishing the objectives and extent of the audit program;
- establishing the responsibilities, resources, and procedures;

- ensuring the implementation of the audit program,
- monitoring and reviewing the audit program to improve its efficiency and effectiveness, and
- ensuring that appropriate program records are maintained.

Because the standard may be applied to internal and external auditing, setting the objectives and extent of the audit program is a critical early step in defining the audit program for a particular organization or application. Any audit program should be managed by persons having appropriate authorities and resources to implement the program.

The audit program may also address the possibility of "combined audits" and "joint audits." A "combined audit" occurs when a QMS and EMS are audited at the same time by the same audit team. A "joint audit" occurs when two audit teams cooperate to audit an organization during the same period with one team auditing the QMS and the other team auditing the EMS.

The audit program should be monitored and reviewed to ensure its ongoing effectiveness in meeting the needs of the organization. Adjustments to the audit program should be made when needed in order foster improvements.

Clause 6 - Audit Activities:

Clause 6 describes the six general steps in planning and conducting an audit. The steps include:

- initiating the audit,
- conducting document review,
- preparing for the on-site audit activities,
- conducting on-site audit activities,
- preparing, approving, and distributing the audit report, and
- completing the audit (including any follow-up activity that may be needed).

Initiating an audit requires consideration of several factors and actions, including:

- appointing a appropriate audit team leader,
- having defined audit objectives,
- confirming that the audit is feasible,
- establishing a satisfactory audit team, and
- establishing the initial contact with the auditee.

Once formed, the audit team will review any available documents pertaining to the audit and prepare for the on-site phase of the audit, including the logistics required and arrangements (such as travel) to be made. Preparation for the on-site audit activities may also include:

- creating an audit plan to document how the audit will be conducted,
- assigning specific work or responsibilities to audit team members, and
- developing work documents such as checklists and sampling plans.

Whether a QMS or EMS audit, the on-site activities are similar and include:

- opening meeting with the auditee,
- roles and responsibilities of guides (as needed),
- collection and verification of information,
- audit findings,
- communication with the audit client and auditee,
- preparation of the closing meeting, and
- closing meeting.

Reporting on the audit results is a critical step and must accurately reflect what transpired during the audit. The key is to address the extent of conformance to the audit criteria, the effectiveness of the management system implementation, and the ability of the management review process to assure the continuing suitability and effectiveness of the management system. This is a significant difference from QMS audit criteria in the past when auditors frequently commented on the suitability and effectiveness of the management system itself. This was inappropriate for two reasons: (1) management is responsible for assessing the value (i.e., "suitability and effectiveness") of the management system and (2) the auditors may lack critical knowledge about the organization's operations in order to assess the value of the management system.

The standard provides for audit follow-up as needed to confirm that all non-conformances have been addressed. In most cases, the audit will be completed when all activities described in the audit plan have been completed; however, there may be occasions when follow-up by the same audit team will be necessary, for example, in an internal audit..

COMPETENCE OF AUDITORS:

Auditors must be competent to perform their assigned tasks and there should be a consistent process for initially selecting and continually evaluating the competence of auditors.

The guidance provides describes the general knowledge, skills, and personal attributes needed for an auditor and an audit team leader. An auditor needs knowledge and skills in audit principles, procedures, and techniques in order to be able to implement the audit. Similarly, the auditor needs to understand the scope of the audit and concepts of management systems in order to apply audit principles effectively. An audit team leader needs to have these same knowledge and skills as well as have the appropriate organizational and leadership skills to be able to

implement the audit consistent with the goals of the audit program. In addition, the auditor and audit team leader will need knowledge and skills pertaining to QMS and their applications and EMS and their applications, as appropriate. When combined audits are required, knowledge and skills in both areas will be necessary.

There is also a need for auditors to have appropriate education, work experience, auditor training, and auditing experience consistent with the needs of the audit program. Typically, the levels of education, training, and experience will vary according to the specific goals and objectives of the audit program. For example, the levels of education, training, and experience needed for internal (first-party) auditors is very likely to differ significantly from those for third-party, certification auditors. In a practical manner, these levels will be set by the "owner" of the audit program or by an appropriate accreditation body. For ISO 19011, there has been a lengthy and yet unresolved debate about what these levels should be and who should set them.

At present, the DIS includes minimum levels of education, training, and experience for third-party certification auditors and recommends their use. Sentiments have been strong among some countries that this is needed "to raise the bar of excellence" for auditors professionally, but representatives of some developing countries have expressed concern that the requirements are too burdensome for them. Clause 7 of the DIS contains a table of "illustration of indicators of education, work experience, auditor training, and audit experience." While ISO 19011 is officially a guideline, inclusion of this table in the standard could be interpreted as meaning that these are minimum levels, and, in fact, makes this recommendation for certification audits in Section 7.6.4. The U.S. believes that this table is inappropriate for this standard and infringes upon the authorities of international and national certification bodies. The table is certainly inconsistent with the ANSI/RAB National Accreditation Programs for registrars for ISO 9001 and ISO 14001. Moreover, the U.S. fears that some users could be influenced to apply the table to other audit situations, including internal audits and second-party supplier audits. The U.S. has proposed that the table be deleted or, as a best case, moved to an Informative Annex of the standard with additional examples that cover the full range and scope of auditing to be addressed by the standard. Each national standards body would be responsible for defining the minimum experience levels appropriate for auditors, recognizing that there are differences between the major industrialized nations and the developing countries in terms of capabilities. In accordance with ISO rules and procedures, the U.S. may vote to disapprove the DIS in order to raise this important issue again with the JWG in Vancouver. The final resolution of this issue is uncertain at present.

The standard includes a process to guide the initial selection of auditors commensurate with the needs of the audit program. Since some audit programs may be long term in nature and auditors may be used over an extended period of time, the standard also describes a process for the on-going evaluation of auditor competence. The maintenance of auditor competence includes continuing professional development, such as through additional training, participation in conferences and

seminars, and additional work experience outside the audit program.

CONCLUSIONS:

The Draft International Standard of ISO 19011 has accomplished several important objectives in the development of a consensus standard:

- the contents of ISO 10011-1, -2, and -3 have been fully incorporated into the standard;
- the contents of ISO 14010, ISO 14011, and ISO 14012 have been fully incorporated into the standard;
- the interests of the environmental and quality communities have been successfully integrated into one document;
- consistency with the requirements and terminology contained in ISO 9001 and ISO 14001 have been accomplished; and
- the new standard has been made easier to use with a logical structure and with a number of diagrams and examples.

While some critical issues remain to be resolved, the U.S. remains optimistic that they will be resolved and that an International Standard will emerge for use by the fall of 2002.

REFERENCES:

1. ISO 19011, *Guidelines on Quality and/or Environmental Management Systems Auditing*. Draft International Standard, International Organization for Standardization, Geneva, Switzerland (May 2001).

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX B: DATA QUALITY ASSESSMENT CHECKLIST

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

Data Quality Assessment Checklist

Partner:

Check-sheet 1 of

Strategic Objective:

Intermediate Result:

Performance indicator:

Data source(s):

Year or period for which the data are being reported:

Date(s) of assessment:

Location(s) of assessment:

Assessment team members:

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:
_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score =]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	❑	❑		
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**1.1.1.**	
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**1.1.2.**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**1.1.3.**	
➢ Were response rates sufficiently large?	❑	❑	**1.1.4.**	
➢ Has non-response rate been followed up?	❑	❑	**1.1.5.**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	❑	❑	**1.1.6.**	
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	❑	**1.1.7.**	
➢ Are definitions for data to be collected operationally precise?	❑	❑	**1.1.8.**	
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**1.1.9.**	
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**1.1.10.**	
Transcription Error				
➢ What is the data transcription process?	❑	❑		
➢ What is the potential for error?	❑	❑		
➢ Are steps being taken to limit transcription error?	❑	❑	**1.1.11.**	
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑		
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator:	**2.**	❑		
➢ Is there manipulation of secondary and/or tertiary data?	**3.**	**4.**		
➢ How are the risks associated with manipulating data identified	**5.**	**6.**		

1. VALIDITY—Do the data adequately represent performance? [Average Score =]				
	Yes	**No**	**Score**	**Comments**
and managed?				
➢ Are the correct formulae being applied?	**7.**	❑	**7.1.1.**	
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**7.1.2.**	
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**7.1.3.**	
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	❑	❑		
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑		
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑		
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑		
➢ Is the sample of adequate size?	❑	❑		
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑		
Non-conformities:				
Strengths and Vulnerabilities:				
Recommendations for improvement (R):				

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	❑	❑		

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =]				
	Yes	**No**	**Score**	**Comments**
➢ Is the same instrument used to collect data from year to year, location to location?	❑	❑		
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	❑	❑		
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑		
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	❑	❑		
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	❑	❑		
➢ Do these procedures provide for periodic sampling and quality assessment of data?	❑	❑	**7.1.4.**	
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	❑		
➢ Are data problems at each level reported to the next level?	❑	❑	**7.1.5.**	
➢ Are data quality problems clearly described in final reports?	❑	❑	**7.1.6.**	
Non-conformities:				
Strengths and Vulnerabilities:				
Recommendations for improvement (R):				

3. TIMELINESS—Are data collected frequently and are they current? [Average score =]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	❑	❑		
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑		

3. TIMELINESS—Are data collected frequently and are they current? [Average score =]				
	Yes	**No**	**Score**	**Comments**
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	❑	❑		
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	❑	❑		
➢ Are the data reported as soon as possible after collection?	❑	❑		
➢ Is the date of collection clearly identified in the report?	❑	❑		
Non-conformities:				
Strengths and Vulnerabilities:				
Recommendations for improvement (R):				

4. PRECISION—Do the data have an acceptable margin of error? [Average score =]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑		
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑		
➢ Have targets been set for the acceptable margin of error?	❑	❑		
➢ Has the margin of error been reported along with the data?	❑	❑		
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	❑		
Non-conformities:				
Strengths and Vulnerabilities:				
Recommendations for improvement (R):				

5. INTEGRITY—Are data free of manipulation? [Average score =]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data	❑	❑		

5. INTEGRITY—Are data free of manipulation? [Average score =]

	Yes	No	Score	Comments
are manipulated for political or personal reasons?				
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	**8.**	❑		
➢ Has there been independent review?	❑	❑		

Strengths and Vulnerabilities:

Non-conformities:

Recommendations for improvement (R):

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score =]

What is the main source of data? (Tick the appropriate box below)

Primary		Secondary		Tertiary		Mixed	

	Yes	No	Score	Comments
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	❑		
➢ Have the risks associated with secondary and tertiary data sources been identified?	❑	❑		
➢ Has the credibility of the secondary and tertiary data been established?	❑	❑		

Strengths and Vulnerabilities:

Non-conformities:

Recommendations for improvement (R):

For indicators for which no recent relevant data are available:

If no recent relevant data are available for this indicator, why not?

What concrete actions are now being undertaken to collect and report this data as soon as possible?

On what date will data be reported?

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX C: INDICATOR QUALITY WORKSHEET

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

INDICATOR QUALITY ASSESSMENT CHECKLIST

Strategic Objective:

Intermediate Result:

Performance indicator:

Is this indicator reported in the R4 Report?

Date(s) of assessment:

Assessment team members:

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

1. DIRECTNESS

Question	Yes	No	Comments
➢ Does the indicator closely measure the result it is intended to measure?	❑	❑	
➢ Does the indicator have credibility, which is defensible in theory and practice?	❑	❑	
➢ Does the indicator represent an acceptable measure to both proponents and skeptics?	❑	❑	

Strengths and Vulnerabilities:

Recommendations for improvement (R):

2. OBJECTIVITY

Question	Yes	No	Comments
➢ Is the indicator unambiguous about what is being measured?	❑	❑	
➢ Is the indicator unambiguous as to what kind of data should be collected?	❑	❑	
➢ Is the indicator unambiguous as to how results should be interpreted?	❑	❑	
➢ Does the indicator measure only one phenomenon at a time?	❑	❑	

Strengths and Vulnerabilities:

Recommendations for improvement (R):

3. PRACTICALITY

Question	Yes	No	Comments
➢ Can data be collected for this indicator easily and reliably?	❑	❑	
➢ Can data be collected for this indicator frequently enough to inform management decisions?	❑	❑	
➢ Are the costs of data collection for this indicator reasonable?	❑	❑	

Strengths and Vulnerabilities:

Recommendations for improvement (R):

4. ADEQUACY

Question	Yes	No	Comments
➢ Is this indicator a measure of progress? (Rather than an attempt to measure everything)	❑	❑	
➢ Does this indicator measure an output? (Rather than an input)	❑	❑	
➢ Does the indicator measure attribution?	❑	❑	

Strengths and Vulnerabilities:

Recommendations for improvement (R):

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)

Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX D: WORK PLAN

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)

Contract No: 674-C-00-01-10051-00

STRATEGIC OBJECTIVE 6 (SO6) AND GLOBAL CLIMATIC CHANGES (GCC) DATA QUALITY ASSESSMENT (DQA)

WORKPLAN (August 11, 2003)

1. OBJECTIVES AND SCOPE OF THE DQA

The primary purpose of this exercise is to assess the quality of data reported by the partners against the SO6 and GCC indicators (FY 2003 PMP). In conducting the DQA, the Assessment Team (Team) will be guided by the criteria described in the ADS 203 (Assessment and Learning) and the PriceWaterhouseCoopers (PWC) 'Performance Management Toolkit'. These criteria include the extent to which the indicators are direct, objective, practical, and adequate. The assessment of data quality will determine the extent to which the data collected by partners for these indicators meet reasonable standards of validity, reliability, timeliness, precision and integrity.

On August 04, 2003, USAID/SA, Khulisa Management Services (Pty) Ltd, Mega-Tech and the Team met and discussed the rationale and context for undertaking the assessment. The parameters of the exercise were clarified and/or confirmed as:

1.1 The primary purpose for conducting the DQA is in order to allow the SO6 Team to assess for any liabilities, contingent or actual, that may arise due to data quality issues and thus implement improvements for the management of risk.

1.2 The DQA must also provide the SO6 Team with a systems analysis so that they will be able to correct and/or improve their own data handling activities.

1.3 The DQA must take into account the time-sensitive nature of the exercise, due to a key staff member being about to depart from USAID/SA.

1.4 The assessment will be limited to those partners listed in Tables 1, 2 and 3.

2. METHODOLOGY

2.1 Sampling:

The list of partners originally supplied to the Team represented an unknown percentage of the total partner Universe reporting against SO6 and GCC. The decision was therefore taken, and agreed to by USAID/SA, that the DQA would be based on Probability-Based Sampling of the given population, understanding that inherent limitations exist with sampling a pre-existing sample. The given list of partners was subdivided into three Populations - those that report data directly to USAID/SA, those that report data through Mega-Tech and those for which data will be reviewed at USAID/SA.

The first Population (Table 1) – those partners reporting directly to USAID/SA - was divided into groups such that the sampling would be representative of the whole. The given list of partners has been selected from the following groups to ensure good representation:

2.1.1	Organizational Size	:	This group is disaggregated by the volume of data reported and thus the degree of risk to USAID/SA.
2.1.2	Organizational Type	:	This group is disaggregated by the legal entity type.
2.1.3	Organizational Location	:	This group is disaggregated by urban location of project(s).
2.1.4	Indicator Spread	:	This group is disaggregated on the number of indicators the partner reports on.

The second Population (Table 2) – those partners reporting through Mega-Tech to USAID/SA – will be subject to a documentation review at the Mega-Tech Offices in Pretoria. All these companies/organizations report data indirectly to USAID/SA through Mega-Tech. The Team will work directly with Mr. Horn of Mega-Tech to obtain the required information such as handling of data and improvements that have been made in terms of data quality.

The third Population (Table 3) incorporates those partners for which data can only be reviewed at USAID/SA, and which have been identified during the process of the DQA for Populations one and two.

Three (3) of the partners on the original list given to the Team, were purposefully excluded from the sampling for the DQA. These partners are DeLoitte and Touche; Deloitte and Touche Emerging Markets Group; and Agrilinks (EM&I). These partners were part of the SO5 DQA.

2.2 DQA Methodology:

Regarding the Team's approach to assessing the quality of data collected and reported on by individual partners, an adapted version of the internationally recognized International Standards Organization (ISO19011) systems auditing approach will be employed. This involves a standard data verification process on site that will be administered by the Team. The approach requires that partners complete the Data Quality Assessment Checklist prior to the on-site visit. All partners must receive the form and must be notified of the requirement to complete it as a matter of urgency.

The Team will then review the information presented in the checklist and perform the verification process of the data on site for Population One and in discussion with Mega-Tech for Population Two. The results of the verification process will allow the Team to assess each organization's capacity to

collect and report on SO6 and GCC indicator data, and will point out strengths and vulnerabilities of the partners' data systems. This information will be contained in the section of the report titled, 'Data Quality Assessment'. Critical to the success of the DQA, in terms of its ability to add value to the Data Quality chain, are the discussions that will be held with various persons from USAID/SA, in particular those that have been and/or are critical in the management of SO6. Undoubtedly various methodological issues will arise in the course of the assessment. The Team will consult with the SO6 team regularly in this regard.

TABLE 1 - POPULATION ONE – REPORT DIRECT TO USAID/SA

PARTNER	AUDITOR	ORG. SIZE	ORG. TYPE	ORG. LOCATION	INDICATOR SPREAD	INDICATORS PER RFP
Ndlandlamuka Local Project (NLP) – Giyani Daniel Mashimbye	JvG	Small	Section 21 NGO	Rural (Municipality but not Metro)	GCC Medium	GCC 2.1
Chemonics (RWD) – Bushbuck Ridge Robert Mbwana	JvG	Medium	Commercial	Rural	SO Narrow GCC Narrow	SO 6.2(a) GCC 4.1
Isandla Partners in Development (IPD) – Port Elizabeth Liesel du Plessis	JvG	Medium	Section 21 NGO	HACD Project - Urban	SO Narrow	PI 6.1.3 PI 6.2.1 PI 6.2.2
Peoples Housing Partnership Trust (PHPT) – Pretoria Pinky Vilakazi	RM	Large	State	Urban	SO Broad	SO 6.1(e) SO 6.2(e) PI 6.1.3 PI 6.2.1 PI 6.2.2
Municipal Infrastructure Investment Unit (MIIU) – Midrand Jackie Lesaoane and Jim Leigland	JvG RM	Large	Section 21 Para-State	Urban With some Rural element	SO Broad GCC Narrow	SO 6.1(a-d) SO 6.2(a-d) PI 6.1.3 PI 6.2.1 PI 6.2.2 GCC 3.5

PARTNER	AUDITOR	ORG. SIZE	ORG. TYPE	ORG. LOCATION	INDICATOR SPREAD	INDICATORS PER RFP
Cato Manor Development Association (CMDA) – Durban Heather Maxwell and Ian Rolf	JvG	Small	Section 21	Urban	GCC Narrow	GCC 3.5
Kwa-Zulu Natal Project Preparation Trust (PPTKN) – Durban Mark Misselhorn	JvG	Small	NGO	Urban/Rural	SO Narrow	PI 6.2.2
Department of Environmental Affairs and Tourism / University of Cape Town (DEAT/UCT) – Cape Town Dr. Merle Sowman	JvG	Small	State - Academic	Urban	GCC Narrow	GCC 1.2
Johannesburg Housing Company (JHC) – Johannesburg Ayesha Rehman	RM	Medium	Section 21	Urban	SO Broad	SO 6.1(e) SO 6.2 (e) PI 6.1.3 PI 6.2.1 PI 6.2.2

TABLE 2 – POPULATION TWO - REPORT VIA MEGA-TECH TO USAID/SA

PARTNER	AUDITOR	INDICATOR SPREAD	REPORTED INDICATORS
Agama Energy (AE)	RM	GCC Broad	GCC 1.2 GCC 3.1 GCC 3.4
Cape Technikon (CT)	DrPR	GCC Narrow	GCC 1.2
Food and Trees for Africa (FTFA)	JvG	GCC Narrow	GCC 2.1
International Institute for Energy Conservation (IIEC)	RM	GCC Broad	GCC 1.2 GCC 3.1 GCC 3.5
Lynedoch Development Foundation (LDF)	DrPR	GCC Narrow	GCC 3.5
Midrand Ecocity (ME)	DrPR	GCC Medium	GCC 1.2 GCC 3.1
National Development Initiative for Social Welfare (NDISWE)	JvG	GCC Medium	GCC 1.2 GCC 3.5
Soweto Development Foundation (SDF)	RM	GCC Narrow	GCC 2.1

TABLE 3 – POPULATION THREE – DATA TO BE REVIEWED AT USAID/SA

PARTNER	AUDITOR	ORG. SIZE	ORG. TYPE	ORG. LOCATION	INDICATOR SPREAD	INDICATORS
First Rand Bank (FRB) – Johannesburg	RM	Large	Commercial	Urban	SO Broad	SO 6.1(a-e) SO 6.2(a-d) IR 6.2.1 IR 6.2.2
Kutlwanong Civic Integrated Housing Trust (KCIHT) – Kimberley	RM	Medium	NGO	Urban/Rural	SO Narrow GCC Narrow	IR 6.2.1 IR 6.2.2 GCC 3.5

2.3 Work Activities:

Given below is a schedule of activities/responsibilities for each Team member (Table 4). The following provides a brief chronological description of workplan activities:

2.3.1 Preliminary review of the contracts of the partners in question (see Tables 1,2 and 3) as well as the other documentation supplied by USAID/SA on August 04, 2003.

2.3.2 On site field audits by Mr. Martin and Mrs. van Graan at the designated partners. Some of the smaller volume-reporting partners will be audited first to ensure that Mr. Martin and Mrs. van Graan are satisfied with the DQA criteria and process, before handling the large volume reporting partners, i.e. First Rand Bank (FRB) and Municipal Infrastructure Investment Unit (MIIU). The Team will prepare indicator quality assessment tables for each relevant indicator, which address the criteria contained in the adapted 'Performance Indicator Quality Assessment'.

2.3.3 Dr. Richards will verify all the DQA's completed by Mr. Martin and Mrs. van Graan. A validation exercise of data quality with each partner based on the information contained in 'Data Quality Assessment Worksheet' using the ISO audit standards method will be conducted simultaneously.

2.3.4 On August 14, 2003 the Team will interview Mr. Kolker of USAID/SA for a history of how the reported data was handled and/or manipulated.

2.3.5 The Team will draft sections of the report in accordance with the time frame contained in Table 4.

TABLE 4 – WORKPLAN CALENDAR FOR SO6 AND GCC DQA

DATE(S)	RESPONSIBLE TEAM MEMBER(S)			ACTIVITY
	RM	DrPR	JvG	
Monday, Aug 04, 2003	✓	✓	✓	Team meeting Khulisa/Mega-tech meeting USAID meeting
Tuesday, Aug 05, 2003	-	-	✓	Develop workplan
Wednesday, Aug 06, 2003	✓	✓	✓	Finalise workplan at Khulisa
Thursday, Aug 07, 2003	✓	✓	✓	Present workplan to USAID/SA and Mega-Tech Documentation preparation for DQA at Khulisa
Friday, Aug 08, 2003	✓	-	-	Documentation preparation for DQA

Saturday, Aug 09, 2003	✓	-	-	Documentation preparation for DQA

DATE(S)	RESPONSIBLE TEAM MEMBER(S)			ACTIVITY
	RM	DrPR	JvG	
Monday, Aug 11, 2003	✓	✓	✓	DQA Mega-Tech (Population two)
Tuesday, Aug 12, 2003	✓	✓-	✓	USAID SO6 Ms Knight (Systems data issues) Documentation preparation for DQA
Wednesday, Aug 13, 2003	✓	-	-	Documentation preparation for DQA
Thursday, Aug 14, 2003	✓	-	✓	USAID/SA Meeting –Mr. Kolker – at Khulisa
Friday, Aug 15, 2003	✓	-	✓	Documentation preparation for DQA
Saturday, Aug 16, 2003	-	-	✓	Documentation preparation for DQA
Monday, Aug 18, 2003	-	-	✓	Fieldwork – NLP
Tuesday, Aug 19, 2003	✓	-	✓	Fieldwork – RWD
Wednesday, Aug 20, 2003	✓	-	✓	Fieldwork – PHPT and IPD
Thursday, Aug 21, 2003	✓	-	✓	Fieldwork – MIIU
Friday, Aug 22, 2003	✓	-	✓	Fieldwork – MIIU
Saturday, Aug 23, 2003	✓	✓	✓	Consolidation and Data Quality Assurance with Dr. Richards
Monday, Aug 25, 2003	✓	-	✓	Fieldwork – CMDA
Tuesday, Aug 26, 2003	✓	-	✓	Fieldwork – PPTKN
Wednesday, Aug 27, 2003	✓	-	✓	Fieldwork – EEU/UCT
Thursday, Aug 28, 2003	✓	-	-	Fieldwork – JHC
Friday, Aug 29, 2003	✓	-	✓	Draft report
Saturday, Aug 30, 2003	-	-	✓	Draft report
Monday, Sept 01, 2003	✓	✓	✓	Draft report
Tuesday, Sept 02, 2003	✓	-	-	Draft report
Thursday, Sept 04, 2003	✓	✓	✓	Draft report
Friday, Sept 05, 2003	✓	-	✓	Touch down with USAID/SA
Saturday, Sept 06, 2003	✓	✓	✓	Draft report
Monday, Sept 08, 2003	✓	✓	-	Draft report
Tuesday, Sept 09, 2003	✓	-	-	Draft report
Wednesday, Sept 10, 2003	✓	-	-	Draft report
Thursday, Sept 11, 2003	✓	✓	✓	Draft report

DATE(S)	RESPONSIBLE TEAM MEMBER(S)			ACTIVITY
	RM	DrPR	JvG	
Friday, Sept 12, 2003	✓	✓	✓	Oral debriefing to USAID/SA and Mega-Tech
Tuesday, Sept 16, 2003	-	-	-	Receive comments from USAID/SA
Wednesday, Sept 17, 2003	✓	-	-	Finalise report
Thursday, Sept 18, 2003	✓	-	-	Finalise report
Friday, Sept 19, 2003				Deliver report to USAID/SA and Mega-Tech
TOTAL NO. OF DAYS ON DQA	**32**	**12**	**26**	**-**

3. REPORT OUTLINE

- Executive Summary
- Background
- Methodology
- Data Quality Assessment
- Indicator Quality Assessment
- Systems Findings
- Recommendations
- Appendices:
  - DQA's Per Partner in Population One
  - DQA's Per Partner in Population Two
  - Performance Indicator Assessment
  - Individuals/Organizations Contacted

4. WORKPLAN ATTACHMENTS

4.1 Indicator Quality Assessment Template

4.2 Data Quality Assessment Template

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX E: DQA NDLANDLAMUKA LOCAL PROJECT

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

NDLANDLAMUKA LOCAL PROJECT (NLP)- GIYANI
Check-sheet 1 of 2

Strategic Objective	**:**	Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	GCC Result 2 – Reduced net greenhouse gas emissions from the land use / forest management sector.
Performance Indicator	**:**	GCC Indicator 2.1 – Area where USAID has initiated interventions to maintain or increase carbon stocks or reduce their rate of loss.
Contractual Obligations	**:**	One (1) SO6 Indicator
Data Source(s)	**:**	Tree Maintenance Contract Agreements; Invoices from Makden Nursery
Year or Period for which the data are being reported	**:**	14 June 2003 (start) to Feb 2003 (completion)
Date(s) of Assessment	**:**	18 August 2003
Location(s) of Assessment	**:**	NLP offices – 2nd Floor Old Mutual Building, Giyani Town, Limpopo Province Daniel Mashimbye, Project Manager Daniel Chauke, Financial Manager
Assessment team members	**:**	On-site was Jacqui van Graan Consolidators – Richard Martin and Dr Penelope Richards

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

Key to Acronyms:
DM : Daniel Mashimbye, Project Manager of NLP
Team : SO6 DQA Team

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.87]					
	Yes	**No**	**Score**	**Comments**	
Face Validity					
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	DM – Yes Team – the number of trees planted is a direct measurement of the activity and the number of trees is what is reported in the GCC data tables obtained from Mega-Tech even though the unit of measure is hectares. The number of trees has not been converted to number of hectares.	
Measurement Error					
***Sampling Error* (only applies when the data source is a survey)**					
➢ Were samples representative?	❑	❑	**N/A**	DM – Yes	
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	Team – sampling for acquisition of trees is 100% representative. All trees bought were counted.	Team – sampling for the counting of actual planted trees could not be determined at audit because the verifiable documentation was not available – these documents were with the person that did the site visits.
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	DM – Yes Team – NOT APPLICABLE – the instrument is not self-reporting.	
➢ Were response rates sufficiently large?	❑	❑	**N/A**		
➢ Has non-response rate been followed up?	❑	❑	**N/A**		
Non Sampling Error					
➢ Is the data collection instrument well designed?	✓	❑	**3**	Team – The "Tree Maintenance Contract Agreement" document (example with Greater Giyani Municipality) clearly states that the donated trees may not be sold for profit and that the recipient will maintain the trees and monitor their growth. The actual number of trees donated to the recipient is written in this agreement. All these agreements were aggregated to obtain the reported result. Also the invoices/delivery notes from the supplier – Makden Nursery – total the number of trees reported. 800 trees were donated to and distributed by the Vembe District to designated schools, etc. (Tax Invoice No. 3210	Team - For actual trees planted, no verifiable data available during audit to check instrument.

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.87]					
	Yes	**No**	**Score**	**Comments**	
				collected 02/09/2003); 411 trees were collected by Mopane District and distributed (Tax Invoice collected 11/09/2003) = 1411 trees as reported in the Grant Activity Completion Report.	
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	DM – No Team – Makden Nursery sold the trees to NLP – verified by invoices. No evidence that invoices are not representative.	Team - For actual trees planted, incentives are unknown as these were not identifiable at audit.
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	DM – Yes Team – The number of trees acquired is the number of trees reported.	
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	DM – No Team – NOT APPLICABLE – there were no enumerators as such. Makden Nursery counted the trees sold and delivered. Makden Nursery is a private business, so some reliance is placed on the integrity of their business practices.	Team – For actual trees planted, this was un-auditable as no information was available.
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**		
Transcription Error					
➢ What is the data transcription process?	✓	❑	**3**	Team – DM writes the report by hand and obtains the data from the invoices/delivery notes. DM also compares this with the trees reported on the "Tree Maintenance Contract Agreements" – these are all totaled. The hand written report is given to the NLP administrator who produces an e-version of the report. DM checks the e-version for mistakes and has any corrections made before sending the report to USAID/SA.	
➢ What is the potential for error?	❑	✓	**3**	Team – LOW transcription error potential. The only potential for error is the aggregation by calculator of the trees on the contracts but because the total is compared to the number of trees actually bought, this error may be picked up quite easily with the comparison.	
➢ Are steps being taken to limit	✓	❑	**3**	Team – DM checks the reports himself	

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.87]					
	Yes	**No**	**Score**	**Comments**	
transcription error?				before sending it to USAID/SA.	
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	Team – NOT APPLICABLE – no transcription data errors were found at audit.	
Data Manipulation					
➢ Do primary data need to be manipulated to produce the data required for the indicator?	❑	✓	**3**	DM – No Team – the primary data is only aggregated – no other calculations or manipulation.	
➢ Is there manipulation of secondary and/or tertiary data?	❑	✓	**3**	DM – No Team – The number of trees acquired, as counted at Makden Nursery, is secondary data and this is only aggregated – no other calculations or manipulation.	
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	Team – NOT APPLICABLE – see above. Also the grant agreement only ran for one year.	
➢ Are the correct formulae being applied?	❑	❑	**N/A**		
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**N/A**		
➢ Have procedures for dealing with missing data been correctly applied?	❑	✓	**2**	DM – No missing data. Team – verification data for the on site visits at tree planting locations was not available. This could therefore not be verified at time of audit. The verification process of counting the actual planted trees was in place as "Memorandum of Goods or Services Required" (e.g. dated 17/04/2003) were seen – these memos are for re-imbursement of monies used for travel/taxi costs to these locations and the reasons are written on the memos. See Non-Conformity 1 below.	
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	DM – Yes Team – Aggregated the invoices/ delivery notes – total correct.	Team – For trees planted – the actual milestone of the contract – the data could not be verified due to the missing data from the person that did the on-site visits.
Representativeness of Data					

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.87]					
	Yes	**No**	**Score**	**Comments**	
➢ Is the sample from which the data are drawn representative of the population served by the activity?	✓	❑	**3**	DM – Yes. Team – the total of the invoice/delivery notes represents the whole (100%) population.	Team – wrt trees actually planted – UNKNOWN - as this could not be verified at audit due to the missing location visits data but the intended sampling appears adequate.
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	✓	**2**	DM – Yes Team – wrt trees actually planted, NO. NLP decided to only visit those locations that were donated more than 10 trees and of these locations not all locations were visited due to financial considerations. As previously stated this could not be verified at audit because of the missing data of these location visits. See Non-Conformity 2 below.	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	✓	❑	**3**	Team – wrt actual trees planted – UNKNOWN – as the information was not available at time of audit. The project is also complete – close out payments have already been made to NLP.	
➢ Is the sample of adequate size?	✓	❑	**3**	DM – Yes Team – wrt trees acquired, YES as 100% population was totaled.	Team – wrt trees planted, UNKNOWN as this could not be verified at audit due to the missing location visits data but the intended sampling framework appears adequate.
➢ Are the data complete? (i.e., have all data points been recorded?)	✓	❑	**3**	DM – Yes Team – wrt trees acquired, all trees bought are allocated and found in the maintenance contracts.	Team – wrt trees planted, NO. Although the location visit data were missing, the intention was met to report complete data.

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.86]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	DM – Yes Team – The data was always transcribed from the data sources - invoices/delivery notes – throughout the project.

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.86]					
	Yes	**No**	**Score**	**Comments**	
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	DM – Yes Team – Data is taken from invoices/ delivery notes. The data source is the instrument. The project also only ran for one year.	
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	❑	❑	**N/A**	DM – Yes Team – NOT APPLICABLE – a single secondary source was used.	
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	Team - NOT APPLICABLE – wrt trees acquired.	WRT trees actually planted – the same system would be used but the project only ran for one year and only one report was handed in.
Internal quality control					
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**3**	DM – Yes Team – Trees counted by Nursery, the delivery notes checked by DM and the trees planted were verified.	
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	DM – Yes Team – Informal procedures. Normal work practice to collect, collate data; also trees planted were verified.	
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	DM – Yes Team – DM checks all data again when reporting to USAID/SA.	
Transparency					
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	✓	❑	**3**	DM – Yes Team – Refer to the Tree Maintenance Contract Agreement.	
➢ Are data problems at each level reported to the next level?	❑	❑	**N/A**	DM – Yes Team – NOT APPLICABLE – DM is the highest level at NLP.	
➢ Are data quality problems clearly described in final reports?	❑	✓	**2**	DM – Yes Team – No mention of data quality is made in the Grant Activity Completion Report. See Non-Conformity 3 below.	

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.50]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	❑	✓	**2**	DM – Yes Team – the project only ran for one year and was only reported once. See Non-Conformity 4 below.
➢ Is a regularized schedule of data collection in place to meet program management needs?	✓	❑	**3**	DM – Yes Team – All the trees were distributed over one week – Arbor Week – therefore no

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.50]				
	Yes	**No**	**Score**	**Comments**
				regular schedule required. Incidental data collection.
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	❑	✓	**2**	DM – Yes Team – The grant was active for one year in which the trees were distributed – reporting at the end of the project. See Non-Conformity 4 below.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	Team – All invoices/delivery notes and tree maintenance contracts fall within the grant agreement period.
➢ Are the data reported as soon as possible after collection?	❑	✓	**2**	DM – Yes Team – the number of trees was only reported annually – the number of trees acquired only reported on Grant Activity Completion Report (GACR). See Non-Conformity 4 below.
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	DM – Yes Team – The GACR refers to the "grant period" but no physical dates mentioned. See Vulnerability 1 below.

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	DM – Yes Team – DM compares the acquired trees from invoices/delivery notes with the Tree Maintenance Contracts. A process to count the actual planted trees was put in place but unfortunately the data required to verify the location visits was still with the person that made these visits – missing data. No actual margin of error calculated but everything is in place for the calculation to be made.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**2**	See Non-Conformity 5 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	✓	**3**	DM – No Team – All the required systems are there but not taken a step further to actually calculate the margin of error.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data	✓	❑	**3**	DM – Yes Team – the trees were counted by a

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
are manipulated for political or personal reasons?				private enterprise subject to GAAP.
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	DM – Yes Team – Refer to the Tree Maintenance Contract.
➢ Has there been independent review?	✓	❑	**3**	DM – Yes - mainly financial data. Team - USAID and Mega-Tech visited NLP offices for DQA – no written record(s) provided to the Team. See Vulnerability 2 below.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3.00]								
What is the main source of data? (Tick the appropriate box below)								
Primary		Secondary		Tertiary		Mixed	✓	
	Yes	**No**	**Score**	**Comments**				
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	DM – Yes Team – data used for reporting was only secondary data – the invoice/delivery notes from Makden Nursery.				
➢ Have the risks associated with secondary and tertiary data sources been identified?	✓	❑	**3**	DM – No Team – Reliant on Makden Nursery operating within GAAP.				
➢ Has the credibility of the secondary and tertiary data been established?	✓	❑	**3**	DM – No Team – Invoices/delivery notes are accepted forms of credibility.				

Non-Conformities:

Non-Conformity 1 — Data Manipulation:
MINOR – There is missing data in the form of the records used to verify the number of trees actually planted at the locations visited. "Memorandum of Goods or Services Required" sheets are evidence that these visits occurred but no record of the visit itself was available – these documents have not been obtained from the person that made the location visits to count the trees actually planted.

Non-Conformity 2 — Representativeness of Data:
MINOR – Not all units of the population had an equal chance of being selected for the sample – visits to locations to count the trees actually planted. Those places where less than 10 trees were donated were not considered for a location visit.

Non-Conformity 3 — Transparency:
MINOR – Data quality problems are not mentioned at all in the Grant Activity Completion Report.

Non-Conformity 4 — Frequency:
MINOR – This project was only for one year and data was only reported once at the end of the grant period. This may result in a possible risk as should there have been problems wrt data quality then it would have been too late for USAID/SA to implement corrective actions related to data quality issues.

Non-Conformity 5 — Precision:

	MINOR – the margin of error has not been determined however only a simple calculation is required as all the systems are there.
Strengths:	
Strength 1	Validity: DM correlated the number of trees bought from Makden Nursery with the total number of trees allocated on the tree maintenance contracts. NLP also made on-site visits of all those recipients that were given more than ten (10) trees – and the trees physically counted. This did produce a problem in that not all could be visited due to resources and the distances involved.
Strength 2	Precision: The pure simplicity of the system makes the margin of error easily calculated.
Strength 3	Integrity: There is little room for inappropriate manipulation of data.
Vulnerabilities:	
Vulnerability 1	Currency: Although the "grant period" is sufficient a time period if referred back to the grant agreement, it would be better to give the actual time period the report refers to, to avoid ambiguity.
Vulnerability 2	Integrity: No written records on previous visits to NLP were seen at audit.
Recommendations for Improvement (R):	
Recommendation 1	Validity: There must be an audit trail to demonstrate the completion of the deliverable, i.e. from the acquired trees to the follow up of the planted trees. There is a gap between the acquisition of trees and the verification of the actual planted trees. The intent is there because there is a process in place to count the actual planted trees but NLP must make sure that they have access to the verification documents. The missing data must be obtained from the person that did the on-site location visits.
Recommendation 2	Transparency: Encourage the partners to report on any issues that may impact on data quality.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

NDLANDLAMUKA LOCAL PROJECT (NLP)- GIYANI
Check-sheet 2 of 2

Strategic Objective	**:**	Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	GCC Result 1 - Increased Participation in the UNFCCC.
Performance Indicator	**:**	GCC Indicator 1.2 - Increased capacity to meet requirements of the UNFCCC, including activities in land use/forestry and energy/industrial/ urban sectors.
Contractual Obligations	**:**	One (1) SO6 Indicator
Data Source(s)	**:**	Attendance registers of training sessions and workshops.
Year or Period for which the data are being reported	**:**	14 June 2003 (start) to Feb 2003 (completion)
Date(s) of Assessment	**:**	18 August 2003
Location(s) of Assessment	**:**	NLP offices – 2nd Floor Old Mutual Building, Giyani Town, Limpopo Province Daniel Mashimbye, Project Manager Daniel Chauke, Financial Manager
Assessment team members	**:**	On-site was Jacqui van Graan Consolidators – Richard Martin and Dr Penelope Richards

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

Key to Acronyms:
DM : Daniel Mashimbye, Project Manager of NLP
Team : SO6 DQA Team

1. VALIDITY—Do the data adequately represent performance? [Average Score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	Provided that attending a workshop/ training activity can be equated to increased capacity. DM counted the actual number of persons capacitated by the NLP workshops and/or training sessions. However **two (2) training activities,** instead of actual number of people, are reported on the GCC data tables obtained from Mega-Tech. See Vulnerability 1 and Vulnerability 2 below.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	Team – NOT APPLICABLE - no survey involved. All workshops/training activities had attendance registers that were filled in and signed by the attendees. Although this occurred, only basic personal detail (e.g. name, signature, etc.) was requested so this primary data source cannot be included into the self-reporting category of a survey questionnaire.
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	Team – the basic personal details and signature of each attendee is sufficient information of their presence at the workshop or training session.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	Team – each attendee had to sign his/her own signature. Handwriting on the attendance registers differed - although not verifiable at audit - but the chances are that the attendees are not fictional.
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	Team – A simple table of name and signature is self explanatory.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Team – Attendees filled in their own details.
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	
Transcription Error				
➢ What is the data transcription process?	✓	❑	**3**	Team – DM writes the report by hand and obtains the data from the attendance registers. DM avoids double counting by physically comparing attendance registers for people that attended more than one workshop/training session. The hand written report is given to the NLP administrator who produces an e-version of the report. DM checks the e-version for mistakes and has any corrections made

1. VALIDITY—Do the data adequately represent performance? [Average Score = 3.00]				
	Yes	**No**	**Score**	**Comments**
				before sending the report to USAID.
➢ What is the potential for error?	❑	✓	**3**	Team – LOW transcription error potential. The only real potential for error is the physical checking for double counting of attendees but even this is re-checked by DM himself.
➢ Are steps being taken to limit transcription error?	✓	❑	**3**	Team – DM checks the reports himself before sending it to USAID/SA.
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	Team – NOT APPLICABLE – no transcription data errors were found at audit.
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator?	❑	✓	**3**	Team – the primary data is only aggregated – no other calculations or manipulation.
➢ Is there manipulation of secondary and/or tertiary data?	❑	❑	**N/A**	Team – NOT APPLICABLE – There is no secondary or tertiary data. Also the grant agreement ran only for one year.
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	
➢ Are the correct formulae being applied?	❑	❑	**N/A**	
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**N/A**	
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	Team – No evidence of missing data was found at audit.
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	Team – Attendees from attendance registers not physically counted as this would have been a lengthy process during audit but it was very clear that the counting was completed with much diligence – intent was clearly shown.
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	✓	❑	**3**	Team – the attendees that were counted represents the whole (100%) population. No sampling – the sample equals the population.
➢ Did all units of the population have an equal chance of being selected for the sample?	✓	❑	**3**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	No sampling involved.
➢ Is the sample of adequate size?	✓	❑	**3**	No sampling – the sample equals the population.
➢ Are the data complete? (i.e., have all data points been recorded?)	✓	❑	**3**	No missing data were found at time of audit. No evidence at audit to suggest otherwise.

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.86]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	Team – The data was always transcribed/ captured from the data source – attendance registers – into the electronic spreadsheet throughout the project.
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	Team – Data was taken from attendance registers. The data source is the instrument. The project also only ran for one year.
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	❑	❑	**N/A**	Team – NOT APPLICABLE – a single primary source was used.
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	Team - NOT APPLICABLE – No sampling involved.
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**3**	Team – Checks are run for errors.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	Team – Informal procedures. Normal work practice to collect and collate data.
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	Team – DM checks all data again when reporting to USAID/SA.
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**3**	Team – no physical document but intent can be proven by the fact that DM checks the attendance registers for double counting.
➢ Are data problems at each level reported to the next level?	❑	❑	**N/A**	Team – NOT APPLICABLE – DM is the highest level at NLP.
➢ Are data quality problems clearly described in final reports?	❑	✓	**2**	Team – No mention of data quality is made in the Grant Activity Completion Report. See Non-Conformity 1 below.

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Team – the project only ran for one year and was reported quarterly. See Vulnerability 3 below.
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	**N/A**	Team – The attendees were counted as a total of all the workshops/training activities therefore no regular schedule required -

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
				incidental data collection.
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	Team – The grant was active for one year in which the workshops/training activities were ran – reporting quarterly and then at the end of the project.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	Team – All workshop/training activity dates fall within the grant agreement period. Attendees were requested to fill in a Workshop Evaluation Form – these are all kept on file.
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	Team – The number of attendees were reported quarterly and then on the Grant Activity Completion Report (GACR).
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	Team – The GACR refers to the "grant period" but no physical dates mentioned. See Vulnerability 4 below.

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	Team – No margin of error calculated.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**2**	See Non-Conformity 2 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	✓	**3**	Crosschecks make for a sound system.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	✓	❑	**3**	Team – The attendance register is filled in and signed by individual people who have nothing to gain from filling the register in untruthfully.
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	
➢ Has there been independent review?	✓	❑	**3**	USAID and Mega-Tech visited NLP offices for DQA – no written record(s) provided to the Team. See Vulnerability 5 below.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3.00]								
What is the main source of data? (Tick the appropriate box below)								
Primary	✓	Secondary		Tertiary		Mixed		

	Yes	No	Score	Comments
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	Team – only primary data is used.
➢ Have the risks associated with secondary and tertiary data sources been identified?	❑	❑	**N/A**	Team – no secondary or tertiary data involved.
➢ Has the credibility of the secondary and tertiary data been established?	❑	❑	**N/A**	

Non-Conformities:	
Non-Conformity 1	Transparency: MINOR – Data quality problems are not mentioned at all in the Grant Activity Completion Report.
Non-Conformity 2	Precision: MINOR – the margin of error has not been determined.
Strengths:	
Strength 1	Integrity: There is little room for inappropriate manipulation of data.
Vulnerabilities:	
Vulnerability 1	Face Validity: The partner reports number of people capacitated whereas the GCC data tables obtained from Mega-Tech is reported as training activities.
Vulnerability 2	Face Validity: ONLY if attending a training activity can be equated to “increased capacity”, can there be a logical relationship between the activity and the indicator. However “increased capacity” refers to the actual increased understanding and way of thinking wrt to the training activity and this has not been measured.
Vulnerability 3	Frequency: This project was only for one year. This may result in a possible risk because should there be problems wrt data quality then it would be too late for USAID/SA to implement corrective actions for “correct” data collection.
Vulnerability 4	Currency: Although the ‘grant period’ is sufficient a time period if referred back to the grant agreement, it would be better to give the actual time period the report refers to, to avoid ambiguity.
Vulnerability 5	Integrity: No written records on previous visits to NLP were seen at audit.
Recommendations for improvement (R):	
Recommendation 1	Reliability: Encourage the partners to report on any issues that may impact on data quality.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX F: DQA CHEMONICS

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

CHEMONICS INTERNATIONAL INC. – RETAIL WATER DISTRIBUTION PROJECT
Check-sheet 1 of 2

Strategic Objective	**:**	Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	RM - (IR 6.4) Improved capacity to apply sustainable participating environmental management principles to local level development. Team -
Performance Indicator	**:**	RM - 6.4.1: Number of low-income communities applying sustainable environment management practices. Team – SO 6.2(a) - Number of households receiving municipal services – water.
Contractual Obligations	**:**	One (1) SO6 Indicator
Data Source(s)	**:**	RM – Project progress reports. Team – Municipalities.
Year or Period for which the data are being reported	**:**	RM – 2002 Team – Year 2002 (facsimile dated 24 October 2002).
Date(s) of Assessment	**:**	19 August 2003
Location(s) of Assessment	**:**	Chemonics offices – Protea Centre, Bushbuck Ridge, Mpumalanga Province Daniel Mashimbye, Project Manager Daniel Chauke, Financial Manager
Assessment team members	**:**	On-site was Jacqui van Graan Consolidators – Richard Martin and Dr Penelope Richards

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

Key to Acronyms:
RM : Robert Mbwana, Chief of Party of RWD
Team : SO6 DQA Team

1. VALIDITY—Do the data adequately represent performance? [Average Score = 1.50]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	❑	✓	**1**	RM – Not directly. Team – RWD is monitoring the water service to the various communities identified in the project. The indicator SO 6.2(a) has a unit of measure of number of households for the water services. See Non-Conformity 1 below.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	✓	**1**	RM – Not applicable. Team – Of all the communities only 5 are proclaimed towns (previous R293 towns) – all the others are rural villages. RM ignored all the villages and focused on the towns only in order to report something to USAID. These towns had the highest level of service with in-house running water so the measurement could be quite easily obtained from the municipalities. In total 135 communities (from Monitoring Plan dated 21 August 2001) are being planned to be serviced with water.
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**1**	RM – Not applicable. Team – UNKNOWN – RM asked each Town Manager the questions in a verbal English interview. RM was satisfied with the answers and based his "extrapolation" on the percentages obtained from the interviews. The lack of an audit trail pertaining to the questions asked during the oral survey represent an absolute risk in terms of audit ability. See Non-Conformity 2 below.
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	RM – Not applicable. Team – NOT APPLICABLE – not self-reporting.
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	❑	✓	**1**	RM – Yes. Team – UNKNOWN – no access to the survey questions at audit. See Non-Conformity 2 below.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	RM – No. Team – No benefit is gained from untruthful information.
➢ Are definitions for data to be collected operationally precise?	❑	✓	**1**	RM – Yes. Team – For PI 6.4.1, the indicator RM thinks he is reporting against - sustainability is an almost immeasurable item. RM understands it as 100% water service 100% of the time period in

1. VALIDITY—Do the data adequately represent performance? [Average Score = 1.50]				
	Yes	**No**	**Score**	**Comments**
				question. See Non-Conformity 3 below.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	RM – Yes - we make use of reports from the service provider operators from DWAF. Team – Secondary data from DWAF.
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	RM – Yes - this information is counter-checked with ward councilors. Team – Secondary data from DWAF.
Transcription Error				
➢ What is the data transcription process?	✓	❑	**3**	RM – Computers. Team – Project Officers transfer from DWAF docs to spreadsheet and RM checks the inputs of the data and formulae – RM checks the entire spreadsheet. A physical check from spreadsheet to reported data is made by RM. Report goes to the Chemonics home office in Washington DC for review before given to USAID/SA. RM also checks the data obtained from sources (DWAF and municipalities) data outliers are queried by RM.
➢ What is the potential for error?	❑	✓	**3**	RM – Minimal. Team – LOW transcription error potential – RM counterchecks everything and the head office also goes through everything before given to USAID.
➢ Are steps being taken to limit transcription error?	✓	❑	**3**	RM – Yes – COP counterchecks reports. Team – RM does a physical check from spreadsheet to reported data.
➢ Have data errors been tracked to their original source and mistakes corrected?	✓	❑	**3**	RM – Yes. Team – Any errors that RM finds is queried and corrected even at the source itself.
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator?	❑	❑	**N/A**	RM – No. Team – No primary data involved.
➢ Is there manipulation of secondary and/or tertiary data?	✓	❑	**1**	RM – Yes. Team – RM "extrapolated" the value of 100% water services for the number of towns based on the time period percentages provided by the Town Managers. This calculation could not be shown at audit. See Non-Conformity 4 below.
➢ How are the risks associated with manipulating data identified and managed?	❑	✓	**1**	Team - RM was "not happy" with the way in which the previous data was reported but it was the best he could do at the time. The Oct 2003 report will be better managed. See Non-Conformity 4 below.
➢ Are the correct formulae being applied?	❑	❑	**1**	RM – Yes. Team – UNKNOWN – RM was not able to

1. VALIDITY—Do the data adequately represent performance? [Average Score = 1.50]				
	Yes	**No**	**Score**	**Comments**
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**1**	show me his "extrapolation". See Non-Conformity 4 below.
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	RM – Yes. Team – No missing data found at audit. Also RM does a physical check from spreadsheet to reported data.
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	❑	✓	**1**	RM – Yes. Team – The number of communities reported (4) in Oct 2002 could not be verified at audit. See Non-Conformity 4 below.
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	✓	**1**	RM – Not applicable. Team – Only 5 communities of 135 were considered for the report. Also these are the proclaimed towns – all rural villages were excluded. See Non-Conformity 5 below.
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	✓	**1**	RM – Not applicable. Team – RM did a "judgement call" and decided to look at the towns only for reporting purposes due to the difficulty of reporting against all the communities at the time. See Non-Conformity 5 below.
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	✓	**1**	RM – Not applicable. Team – Only 5 of 135 communities were considered – not comprehensive. See Non-Conformity 5 below.
➢ Is the sample of adequate size?	❑	✓	**1**	RM – Not applicable. Team - Only 5 of 135 communities were considered. See Non-Conformity 5 below.
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	✓	**1**	RM – Not applicable. Team – No rural communities were considered for reporting purposes. See Non-Conformity 5 below.

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.89]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	RM – Yes. Team – Data is transcribed directly from DWAF and the municipalities.
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	RM – Yes. Team – Data is obtained from DWAF and the municipalities – the data source is the

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.89]				
	Yes	**No**	**Score**	**Comments**
				instrument. Oct 2003 will only be the second annual report.
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**3**	RM – Yes. Team – The data obtained from both DWAF and the municipalities are captured and manipulated in the same way. Municipalities had aggregated bank accounts – water is now in a separate bank account after RM requested it – will be separate for Oct 2003 report.
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	RM – Not applicable. Team – This could not be audited because only one report has been handed in as yet – also they only report annually. See Vulnerability 1 below.
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**3**	RM – Yes - ???? Team – RM physically crosschecks everything himself.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	RM – Yes. Team – RM physically crosschecks everything himself and any outliers are queried immediately with the party responsible for the data capture.
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	RM – No. Team – RM checks all data again before reporting to USAID/SA.
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**3**	RM – No. Team – No written procedure but unwritten procedure in place – intent can be proven.
➢ Are data problems at each level reported to the next level?	✓	❑	**3**	RM – Yes. Team – Reported to supervisor in Washington DC. No evidence was seen at audit but RM said he may be able to find these on e-mail record. See Vulnerability 2 below.
➢ Are data quality problems clearly described in final reports?	❑	✓	**2**	RM – Yes. Team – No mention is made of data problems at all. See Non-Conformity 6 below.

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	❑	✓	**2**	RM – Yes. Team – RWD only reports on an annual basis. See Non-Conformity 7 below. Municipalities should be able to report to

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
				RWD on a monthly basis or short notice if required because everything is on their records.
➢ Is a regularized schedule of data collection in place to meet program management needs?	✓	❑	**3**	RM – Yes. Team – RM has requested municipalities to report quarterly to RWD. RM did state that this does not always happen and he has to ask where his data is especially when USAID asked for a report on their sometimes irregular basis.
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	❑	✓	**2**	RM – Yes. Team – RWD was to report only annually but shorter periods are possible. See Non-Conformity 7 below.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	RM – Yes. Team – The contract started in September 2002 and is still ongoing.
➢ Are the data reported as soon as possible after collection?	❑	✓	**2**	RM – Yes. Team – NOT APPLICABLE – reporting only annually. See Non-Conformity 7 below.
➢ Is the date of collection clearly identified in the report?	❑	✓	**2**	RM – No - this is not required in the standard reporting form. Team – the standard report format from USAID has no allocated space for the date of collection. See Non-Conformity 8 below.

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 1.60]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	RM – Not certain this is applicable. Team - No margin of error calculated.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**2**	See Non-Conformity 9 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	✓	**1**	RM – Not certain this is applicable. Team – See Non-Conformities 2 through 5 below.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	✓	❑	**3**	RM – Yes - the data is provided by operators and counter-checked by ward councilors. Community members would provide checks. Team – Unwritten mechanisms would prevent this because the communities would always be there to confer or argue any point made.
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	RM – Yes. Team – Data is obtained from municipalities.
➢ Has there been independent review?	✓	❑	**3**	RM – Yes – as part of the overall project review. Team – USAID did a DQA on 07 Oct 2002 however the content of the DQA report is in question.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3.33]				
What is the main source of data? (Tick the appropriate box below)				
Primary		Secondary	✓	Tertiary		Mixed	
	Yes	**No**	**Score**	**Comments**
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	RM – No - not certain if applicable. Team – Only Secondary data involved.
➢ Have the risks associated with secondary and tertiary data sources been identified?	❑	✓	**2**	RM – No - not certain if applicable. Team – Risks of using secondary data from municipalities not verified. See Non-Conformity 10 below.
➢ Has the credibility of the secondary and tertiary data been established?	❑	✓	**2**	

Non-Conformities:

Non-Conformity 1 — Face Validity:
MAJOR – Although water services is measured in both the indicator and by RWD, the absolute risk is introduced by the fact that the indicator measures number of households but RWD reports number of communities – the numbers will be completely different to what is expected by USAID/SA.

Non-Conformity 2 — Sampling Error:
MAJOR – RWD could not show the questions asked at the oral survey so the survey questions could not be audited.

Non-Conformity 3 — Non Sampling Error:
MAJOR – RM has an elaborate spreadsheet in an attempt to determine sustainability and actual population. The definitions are not operationally precise.

Non-Conformity 4 — Data Manipulation:
MAJOR – RM verbally described an elaborate 'extrapolation' as to how the eventual 4 communities were reported. This could not be audited at all because neither the survey questions nor the calculation used could be shown for audit – this introduces

	an absolute risk.
Non-Conformity 5	Representativeness of Data: MAJOR – Only 5 of a total of 135 communities were used for reporting purposes – only 3.7% of the total population. This is an absolute risk. Besides the small population percentage, only the proclaimed towns (the 5 communities) were considered for reporting purposes – the other communities are all rural (previous R293) villages.
Non-Conformity 6	Transparency: MINOR – Data quality problems are not mentioned in the report.
Non-Conformity 7	Frequency: MINOR – The data has only been reported once so far and reporting is only done annually. This may not be frequent enough to inform program management decisions.
Non-Conformity 8	Currency: MINOR – No date of collection is identified in the report.
Non-Conformity 9	Precision: MINOR – The margin of error has not been determined.
Non-Conformity 10	Data Source Type: MINOR – The possible risks associated with municipalities as data sources have has not been identified.
Strengths: Nil.	
Vulnerabilities:	
Vulnerability 1	Consistency: Only one report has been handed in so the consistency of the sampling method could not be audited. USAID/SA should be aware that should the same sampling method be used in the 2003 report then there will be the same risks highlighted in this DQA. It must be noted that RWD have changed all their processes and will apparently be reporting on all the communities for 2003.
Vulnerability 2	Transparency: RM was not able to show the e-mail records to the supervisor without having to do a major search in his 'Inbox" – the records were not insisted on.
Recommendations for improvement (R):	
Recommendation 1	Validity: Encourage the partners to ensure they have a verifiable audit trail for all data they report to USAID/SA.
Recommendation 2	Reliability: Encourage the partners to report on any issues that may impact on data quality.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

CHEMONICS INTERNATIONAL INC. – RETAIL WATER DISTRIBUTION PROJECT
Check-sheet 2 of 2

Strategic Objective	**:**	Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	-
Performance Indicator	**:**	SO 6.1(a) - Rand value of municipal services – water NB – Attachment 9 is a list of monies available to the project as a whole and not monies spent on water services only. RM intends reporting this as a percentage – money in from the grant (which theoretically should be the same value for expected monies) in relation to actual monies collected/paid for services and thus returned into the system.
Contractual Obligations	**:**	One (1) SO6 Indicator
Data Source(s)	**:**	Business Plans and official letters
Year or Period for which the data are being reported	**:**	20 Sep 2000 to 30 Sep 2002
Date(s) of Assessment	**:**	19 August 2003
Location(s) of Assessment	**:**	Chemonics offices – Protea Centre, Bushbuck Ridge, Mpumalanga Province Daniel Mashimbye, Project Manager Daniel Chauke, Financial Manager
Assessment team members	**:**	On-site was Jacqui van Graan Consolidators – Richard Martin and Dr Penelope Richards

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.93]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	❑	✓	**2**	RWD is reporting the total Rand value of the project as a whole from contracts and official letters. The indicator SO 6.1(a) is for the Rand value of water services ONLY. See Non-Conformity 1 below.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	No survey done.
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	Contracts and official letters are an accepted data source for Rand values.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	Working with money that belongs to sponsors. These sponsors all have financial audits against GAAP.
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	Rand value of monies donated needs no further explanation.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Data obtained from contracts and official letters – no enumerators at RWD level.
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	
Transcription Error				
➢ What is the data transcription process?	✓	❑	**3**	RM transcribes Rand values from contracts and offical letters into the report and checks the inputs of the data himself. Report goes to the Chemonics home office in Washington DC for review before given to USAID/SA.
➢ Is there a potential for error?	❑	✓	**3**	LOW transcription error potential – checks are made to ensure correct data transfer.
➢ Are steps being taken to limit transcription error?	✓	❑	**3**	RM crosschecks the data inputs himself and the head office sees the report before it goes to USAID/SA.
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	No data errors have been found to date. At audit two values reported on Attachment 9 for the period 20/09/2000 to 30/09/2002 could not be verified due to documents not on the RWD premises. RM did try and obtain these docs during the audit but was unsuccessful.
Data Manipulation				

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.93]				
	Yes	**No**	**Score**	**Comments**
➢ Do primary data need to be manipulated to produce the data required for the indicator?	❑	❑	**N/A**	Only secondary data involved.
➢ Is there manipulation of secondary and/or tertiary data?	❑	✓	**3**	Only simple aggregation was required.
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	No manipulation of data.
➢ Are the correct formulae being applied?	❑	❑	**N/A**	
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**N/A**	
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	No evidence of missing data found at audit.
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	Aggregations add up correctly.
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	✓	❑	**3**	No sampling – sample equals the population.
➢ Did all units of the population have an equal chance of being selected for the sample?	✓	❑	**3**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	✓	❑	**3**	No sampling involved.
➢ Is the sample of adequate size?	✓	❑	**3**	No sampling – sample equals the population.
➢ Are the data complete? (i.e., have all data points been recorded?)	✓	❑	**3**	No evidence of missing data found at time of audit.

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.67]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	Data is transcribed directly from contracts and official letters. Also they have only reported once on this data.
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	Data is transcribed directly from contracts and official letters – the data source is the instrument. Also they have only reported

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.67]				
	Yes	**No**	**Score**	**Comments**
				once on this data.
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**3**	Transcribed from what has been deemed to be reliable sources – contracts and official letters.
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	No sampling.
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**3**	RM checks himself and the head office sees the report before it goes to USAID/SA.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	❑	✓	**2**	Checks are only made when capturing the data. See Non-Conformity 2 below.
➢ Do these procedures provide for periodic sampling and quality assessment of data?	❑	✓	**2**	No procedures in place for periodic review. See Non-Conformity 2 below.
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**3**	Informal procedures. No cleaning applicable as data obtained directly from contracts and official letters.
➢ Are data problems at each level reported to the next level?	✓	❑	**3**	Reported to supervisor in Washington DC. No evidence was seen at audit but RM said he might be able to find these on e-mail record.
➢ Are data quality problems clearly described in final reports?	❑	✓	**2**	No mention is made of data problems at all. See Non-Conformity 3 below.

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.40]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	❑	✓	**2**	RWD only reports on an annual basis. See Non-Conformity 4 below.
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	**N/A**	Since data is obtained from contracts and official letters, data collection is incidental.
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	❑	✓	**2**	RWD only reports on an annual basis on this indicator but he could report more often if required. See Non-Conformity 4 below.
➢ Are data from within the policy	✓	❑	**3**	All data fall after the start date of the

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.40]				
	Yes	**No**	**Score**	**Comments**
period of interest? (i.e., are data from a point in time after intervention has begun?)				contract.
➢ Are the data reported as soon as possible after collection?	❑	✓	**2**	RWD only reports on an annual basis. See Non-Conformity 4 below.
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	Attachment 9 clearly indicates the report period.

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	No margin of error calculated.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**2**	See Non-Conformity 5 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	✓	❑	**3**	There are sufficient data checks.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	✓	❑	**3**	All Rand values are taken directly from contracts or official letters – manipulation is unlikely.
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	Rand values on contracts and official letters are determined by the donators.
➢ Has there been independent review?	✓	❑	**3**	USAID did a DQA on 07 Oct 2002 however the content of the DQA report is in question.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3.00]							
What is the main source of data? (Tick the appropriate box below)							
Primary		Secondary	✓	Tertiary		Mixed	

	Yes	**No**	**Score**	**Comments**
➢ Is there any aggregation of different data types into a single	❑	✓	**3**	Only secondary data involved.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3.00]				
figure for reporting purposes?				
➢ Have the risks associated with secondary and tertiary data sources been identified?	✓	❑	**3**	Risk is minimal because Rand values on contracts and official letters pre-determined.
➢ Has the credibility of the secondary and tertiary data been established?	✓	❑	**3**	Contracts and official letters are accepted credible sources of Rand value data.

Non-Conformities:	
Non-Conformity 1	Face Validity: MINOR – RWD is reporting total leveraged funds for the project as a whole not for water services only, which is what the indicator description requires.
Non-Conformity 2	Internal Quality Control: MINOR – There are no procedures for periodic review of data collection or maintenance.
Non-Conformity 3	Transparency: MINOR – Data quality problems are not mentioned in the report.
Non-Conformity 4	Timeliness: MINOR – The data has only been reported once so far and reporting is only done annually. This may not be frequent enough to inform program management decisions.
Non-Conformity 5	Precision: MINOR – The margin of error has not been determined.
Strengths: Nil.	
Vulnerabilities: Nil.	
Recommendations for improvement (R):	
Recommendation 1	Timeliness: Encourage the partners to report on a more frequent basis – not just annually.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX G: DQA ISANDLA PARTNERS IN DEVELOPMENT

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

ISANDLA PARTNERS IN DEVELOPMENT (IPD)
Check-sheet 1 of 1

Strategic Objective	**:**	No. 6 - Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	IR 6.2 - Previously ineligible households developers, builders and municipal service providers obtaining access to credit
Performance Indicator	**:**	IR 6.2.1 - Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision. **AND** IR 6.2.2 - Number of households assisted to obtain shelter/urban services through the provision of credit and subsidies to low income communities.
Contractual Obligations	**:**	1. Number of historically disadvantaged households assisted to obtain shelter and services. 2. Amount of credit or subsidies leveraged. 3. Emission of carbon dioxide equivalents avoided.
Data Source(s)	**:**	Subsidy approvals from PHB.
Year or Period for which the data are being reported	**:**	June 2001 to February 2003
Date(s) of Assessment	**:**	20 August 2003
Location(s) of Assessment	**:**	IPD offices – Pier 14, 444 Govan Mbeki Street, Port Elizabeth Liesel du Plessis, COP Phil Goduka, Executive Director
Assessment team members	**:**	On-site was Jacqui van Graan Consolidators – Richard Martin and Dr Penelope Richards

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:
_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.93]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	IPD reported number of households and Rand value for shelter and services. There is a direct relationship to that which had to be measured.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	NOT APPLICABLE – No survey involved.
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	Database printout of Provincial Housing Board (PHB) of the approved subsidies by ID number and erf number. This is a form designed by PHB/gov.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	This is an approval system based on criteria – the criteria for subsidy are met or they are not met.
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	An approved subsidy represents a house (by the ID number of beneficiary) and the Rand value because the Rand value is R18 400 for each house.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Subsidies are approved by PHB – secondary data.
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	
Transcription Error				
➢ What is the data transcription process?	✓	❑	**3**	COP checks the approved subsidy database obtained from PHB for obvious outliers (all the subsidies are the same Rand value). The Project Secretary captures all data through the entire implementation process. The COP does "tracking" when she gets a copy of the "erf spreadsheet" – looks for obvious outliers and discrepancies. The spreadsheet is printed on a regular basis and COP checks it again for outliers.
➢ Is there a potential for error?	❑	✓	**3**	LOW transcription error potential – double crosschecks are made.
➢ Are steps being taken to limit transcription error?	✓	❑	**3**	Double crosschecks are made – first after capture then regularly thereafter.
➢ Have data errors been tracked to	✓	❑	**3**	IPD found mistakes on the data source

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.93]					
	Yes	**No**	**Score**	**Comments**	
their original source and mistakes corrected?				and had these mistakes corrected at PHB.	
Data Manipulation					
➢ Do primary data need to be manipulated to produce the data required for the indicator?	❑	❑	**N/A**	NOT APPLICABLE - Only secondary data involved.	
➢ Is there manipulation of secondary and/or tertiary data?	❑	✓	**3**	Data obtained directly from data source with a simple aggregation.	
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	NOT APPLICABLE – no data manipulation.	
➢ Are the correct formulae being applied?	❑	❑	**N/A**		
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**N/A**		
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	No evidence of missing data was found at audit – the double crosscheck system of IPD eliminates the possibility of missing data.	
➢ Are final numbers reported accurate? (e.g. does a number reported as a "total" actually add up?)	❑	✓	**2**	Some "total to date" aggregations are incorrect on the table used for reporting purposes. See Non-Conformity 1 below.	
Representativeness of Data					
➢ Is the sample from which the data are drawn representative of the population served by the activity?	✓	❑	**3**	For Tjoks as everybody in the geographical area of Tjoks could apply for a subsidy – sample is equal to the population.	For Motherwell – because the option to apply for subsidy was exclusive to Sathezethu Ngomanyano Housing Association (SNHA) (voluntary) membership.
➢ Did all units of the population have an equal chance of being selected for the sample?	✓	❑	**3**		
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	No sampling involved.	
➢ Is the sample of adequate size?	✓	❑	**3**	For Tjoks as everybody in the geographical area of Tjoks could apply for a subsidy – sample is equal to the population.	Motherwell: The option to apply for subsidy was exclusive to SNHA (voluntary) membership.
➢ Are the data complete? (i.e., have all data points been recorded?)	✓	❑	**3**	No missing data points could be found at audit. The double crosscheck process eliminates errors.	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 3.00]					
	Yes	**No**	**Score**	**Comments**	
Consistency					
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	PHB subsidy approval data have been used the entire period of the grant as well as "Completion of Works" and "Handover Certificates" for number of houses completed.	
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	PHB subsidy approval data have been used the entire period of the grant and "Handover Certificates" for number houses completed. The data source is the instrument.	
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**3**	All data sources are official documents from a single source – one document equals one house.	
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**3**	Tjoks – No sampling involved.	Motherwell – One had to be a member of SNHA.
Internal quality control					
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**3**	Double crosschecks are made – first after capture then regularly thereafter.	
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	Double crosschecks are made – first after capture then regularly thereafter. Should a Ward Councillor agree to a swop of house ownership, IDP would not accept this until signatures were obtained from both parties before they corrected the details on the spreadsheet.	
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	Data provided by PHB are also checked for errors against the initial subsidy applications handed in.	
Transparency					
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	✓	❑	**3**	Sales Process, Tjoks Construction-Administration Process and Chart Flow show all the steps involved in the actual implementation as well as data capture involved in the project as a whole.	
➢ Are data problems at each level reported to the next level?	✓	❑	**3**	Liesel du Plessis reports data problems to the Executive Directive but she is responsible for the follow up of outliers or other data problems for the housing projects. No data quality problems have been recorded to date.	
➢ Are data quality problems clearly described in final reports?	✓	❑	**3**	A separate section on the graphs they send to USAID/SA covers problems/ delays	

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Data are reported each time a deliverable is completed but also on an annual basis. A good system and database allows reporting to be done at any time.
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	**N/A**	The data collection of Rand value and houses subsidised is obtained incidentally on a continuous basis as approvals come through from PHB. However Sales Process, Tjoks Construction-Administration Process and Chart Flow show all the steps involved in the actual implementation as well as data capture involved in the project as a whole. Most items are updated on a weekly basis.
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	Quarterly and bi-annual reports are handed to USAID/SA.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	All reports dated within the grant period.
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	Quarterly and bi-annual reports are handed to USAID/SA.
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	A section for "Report Date" has been set-aside on the table for date of collection, e.g. Sep 2001 – Sep 2002.

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	No margin of error has been calculated.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**2**	No margin of error calculated therefore no target. See Non-Conformity 2 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	✓	❑	**3**	As little error exists in current system.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	✓	❑	**3**	Physical official documents are required for each data point
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	
➢ Has there been independent review?	✓	❑	**3**	USAID/SA completed a DQA on 06 January 2003. An evaluation was completed by United Sector Network (USN) (Sue Marshall) as a local government project.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3]								
What is the main source of data? (Tick the appropriate box below)								
Primary		Secondary		✓	Tertiary		Mixed	
	Yes	**No**	**Score**	**Comments**				
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	Only secondary data involved.				
➢ Have the risks associated with secondary and tertiary data sources been identified?	✓	❑	**3**	IPD check the data from PHB against the original subsidy application information they submitted.				
➢ Has the credibility of the secondary and tertiary data been established?	✓	❑	**3**	Subsidy approvals from PHB are accepted credible data sources.				

Non-Conformities:	
Non-Conformity 1	Data Manipulation: MINOR – Some aggregations did not total correctly on the report.
Non-Conformity 2	Precision: MINOR – The margin of error has not been determined.
Strengths: Nil.	
Vulnerabilities: Nil.	
Recommendations for improvement (R): Nil.	

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX H: DQA PEOPLES HOUSING PARTNERSHIP TRUST

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

Partner: People's Housing Process Trust
Check-sheet 1 of 2

Strategic Objective: SO6: Increased access to shelter and environmentally sound municipal services

Intermediate Result: IR 6.1.3: Number of shelter units completed

Performance indicator: 2 (e) Number of households receiving services – housing

Contractual Obligations: Annual report on number of units completed

Data source(s): Annual Report 2003

Year or period for which the data are being reported: April 31 2002 – March 31 2003

Date(s) of assessment: 20 August 2003

Location(s) of assessment: Offices of the PHPT, Pretoria

Assessment team members: Richard Martin

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:
_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	Number of shelter units produced is the primary indicator of the program, which is a component of the national housing delivery system. However, there is a major question of attribution, i.e. to what extent the housing being developed can be attributed to the PHPT and, by inference, to USAID.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	Not applicable
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	Use of the provincial data bases
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Data is collected via project management systems – no special enumerators involved
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	Not applicable
Transcription Error				
➢ What is the data transcription process? (low risk?)	✓	❑	**3**	Data is aggregation of accounting records by computer: effectively no transcription used
➢ What is the potential for error (low risk?)?	✓	❑	**3**	Zero
➢ Are steps being taken to limit transcription error?	❑	❑	**N/A**	Not applicable
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	Not applicable
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator:	❑	✓	**3**	Not applicable
➢ Is there manipulation of	❑	✓	**3**	Not applicable

1. VALIDITY—Do the data adequately represent performance? [Average Score = 3.00]				
	Yes	**No**	**Score**	**Comments**
secondary and/or tertiary data?				
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	Not applicable
➢ Are the correct formulae being applied?	❑	❑	**N/A**	Not applicable
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	✓	❑	**3**	The same data collection protocols are applied.
➢ Have procedures for dealing with missing data been correctly applied?	✓	❑	**3**	Data cleaning is undertaken before the data is submitted for reporting by the Provinces
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	No evidence at audit to suggest otherwise.
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑	**N/A**	Not applicable – no sampling
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑	**N/A**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	
➢ Is the sample of adequate size?	❑	❑	**N/A**	
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑	**N/A**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.89]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	Consistent data collection, using same method.
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.89]				
	Yes	**No**	**Score**	**Comments**
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**3**	
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	Not applicable – no sampling
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	❑	❑	**3**	The nature of the data and the fact that it is 100% sample means that there is no risk of bias.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	Data cleaning is undertaken before subsidies are granted
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**2**	The process for submitting applications for subsidies and approval are in writing; the data bases is standardized, but quality assessment procedures are not documented in writing (See non-conformity 1)
➢ Are data problems at each level reported to the next level?	❑	✓	**3**	They are dealt with at the Provincial level
➢ Are data quality problems clearly described in final reports?	❑	✓	**3**	Not necessary as they resolved at the point of entry

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Data are available monthly, but are only reported to USAID when required
➢ Is a regularized schedule of data collection in place to meet program management needs?	✓	❑	**3**	
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	
➢ Are the data reported as soon as	✓	❑	**3**	The reports are available from a data

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
possible after collection?				base, but transmission of the data to USAID at the official level is annual
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 3.00]

	Yes	No	Score	Comments
➢ Is the margin of error less than the expected change being measured?	✓	❑	**3**	Primary data used, so no margin of error permitted
➢ Is the margin of error acceptable given the likely management decisions to be affected?	✓	❑	**3**	
➢ Have targets been set for the acceptable margin of error?	❑	❑		Not applicable
➢ Has the margin of error been reported along with the data?	❑	❑		Not applicable
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	❑		Not applicable

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]

	Yes	No	Score	Comments
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	❑	❑		Not applicable
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	
➢ Has there been independent review?	✓	❑	**3**	

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3]

What is the main source of data? (Tick the appropriate box below)

Primary	✓	Secondary		Tertiary		Mixed	

	Yes	No	Score	Comments
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	
➢ Have the risks associated with secondary and tertiary data sources been identified?	✓	❑	**3**	Secondary data sources are cross referenced with other sources and discrepancies become apparent easily
➢ Has the credibility of the secondary and tertiary data been established?	✓	❑	**3**	Standardized national reporting system is used

Non-conformities:

Non-conformity 1: Transparency
MINOR - Data collection, cleaning, analysis, reporting, and quality assessment procedures are not documented in writing

Strengths:

Strength 1: Validity
The unit of measurement is well defined and represents the focus of the program

Strength 2: Validity
The data is obtained from well audited primary sources

Strength 3: Reliability
The data are part of a national, well-audited data base which has been in use for ten years

Strength 4: Timeliness
The data is produced from a live data base managed by each Province, and can be accessed at any time

Strength 5: Precision
No sampling is involved

Strength 6: Integrity
Data is cross checked against auditable facts

Strength 7: Data Source Type
Well tested and audited data is used

Vulnerabilities:

Vulnerability 1: Validity
Attribution of all housing developed through the "People's Housing Process" to the PHPT, and by inference to USAID raises some difficulties, especially in light of the very large numbers reported

Vulnerability 2: Reliability
The data cleaning and checking process is undertaken by a third party – the Provinces – and cannot be verified.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

Partner: People's Housing Process Trust
Check-sheet 2 of 2

Strategic Objective: SO6: Increased access to shelter and environmentally sound municipal services

Intermediate Result: 6.2.1 Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision

Performance indicator: 6.1 (e) Rand Value of municipal services completed - housing

Contractual Obligations: Annual reporting in terms of number of units completed

Data source(s): Annual Report 2003

Year or period for which the data are being reported: April 31 2002 – March 31 2003

Date(s) of assessment: 20 August 2003

Location(s) of assessment: Offices of the PHPT, Pretoria

Assessment team members: Richard Martin

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:
_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.62]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	The number of shelter units produced is the primary indicator of the program which is a component of the national housing delivery system: Rand value is derived from that
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	Not applicable
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	The partner reports in terms of the number of units completed, that instrument is well designed
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	
➢ Are definitions for data to be collected operationally precise?	✓	❑	**2**	Assumptions are made about the Rand value per subsidy in order to arrive at the data used. See non-conformity 1.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Data is collected via project management systems – no special enumerators involved
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	Not applicable
Transcription Error				
➢ What is the data transcription process?	❑	❑	**3**	Primary data is aggregation of accounting records by computer: effectively no transcription used
➢ What is the potential for error?	❑	❑	**3**	In terms of transcription of primary data, zero
➢ Are steps being taken to limit transcription error?	❑	❑	**N/A**	Not applicable
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	Not applicable
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator:	✓	❑	**2**	See non-conformity 1.
➢ Is there manipulation of secondary and/or tertiary data?	✓	❑	**2**	

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.62]				
	Yes	**No**	**Score**	**Comments**
➢ How are the risks associated with manipulating data identified and managed?	❑	✓	**2**	No specific risks are identified, but errors may occur due to selection of incorrect subsidy amount. See non-conformity 1
➢ Are the correct formulae being applied?	✓	❑	**2**	There are risks of over- and under-counting due to the assumption of a standardized subsidy amount. In practice the amount will vary above and below the amount used. On balance there is probably under-reporting of the Rand values. See non-conformity 1
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	✓	❑	**3**	The same data collection protocols are applied.
➢ Have procedures for dealing with missing data been correctly applied?	✓	❑	**3**	Data cleaning is undertaken before the data is submitted for reporting by the Provinces
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑	**N/A**	Not applicable – no sampling
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑	**N/A**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	
➢ Is the sample of adequate size?	❑	❑	**N/A**	
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑	**N/A**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	Yes, the national data base system
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**N/A**	
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	Not applicable – no sampling
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	❑	❑	**3**	The nature of the data and the fact that it is 100% sample means that there is no risk of bias.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	Data cleaning is undertaken before subsidies are granted
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**3**	The process for submitting applications for subsidies and approval are in writing; the data bases is standardizes, but quality assessment procedures are not documented in writing
➢ Are data problems at each level reported to the next level?	❑	✓	**N/A**	They are dealt with at the Provincial level
➢ Are data quality problems clearly described in final reports?	❑	✓	**N/A**	Not necessary as they resolved at the point of entry

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Data are available monthly, but are only reported to USAID when required
➢ Is a regularized schedule of data collection in place to meet program management needs?	✓	❑	**3**	
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]

	Yes	No	Score	Comments
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.40]

	Yes	No	Score	Comments
➢ Is the margin of error less than the expected change being measured?	✓	❑	**2**	The margin of error is almost certainly very small due to the nature of the data, but information to allow checking is not available. See non-conformity 2
➢ Is the margin of error acceptable given the likely management decisions to be affected?	✓	❑	**3**	
➢ Have targets been set for the acceptable margin of error?	❑	❑	**2**	No. See non-conformity 2
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	The calculations on which data are based are approximate, but the margin of error has not been stated. See non-conformity 2
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	✓	❑	**3**	A greater degree of accuracy would only be possible with significantly more complex reporting procedures. It is unlikely that these would reveal substantially different data.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]

	Yes	No	Score	Comments
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	❑	❑	**N/A**	Not applicable
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	Objectivity of the data is ensured by financial and other audits
➢ Has there been independent review?	✓	❑	**3**	Financial and performance audits are undertaken

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 2.33]

What is the main source of data? (Tick the appropriate box below)

Primary		Secondary	✓	Tertiary		Mixed	

	Yes	No	Score	Comments
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	No
➢ Have the risks associated with	❑	✓	**2**	The use of standardized subsidy amounts

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 2.33]				
secondary and tertiary data sources been identified?				poses a risk
➢ Has the credibility of the secondary and tertiary data been established?	✓	❑	**3**	The subsidy amounts are regulated by national government, so the credibility cannot be questioned

Non-conformities:
Non-conformity 1: Validity
MINOR - The data for reporting on this is the number of units: the partner does not report on Rand Value. In practice the range of subsidies provided is very small, so this vulnerability is minor

Non-conformity 2: Precision
MINOR – Although no margin of error has been established, the likely margin is less than the change been affected by the project.

Strengths:
Strength 1: Reliability
The primary data collection has followed a consistent and easily auditable process

Strength 2: Timeliness
Data are taken from a live data base

Vulnerabilities:
Vulnerability 2: Precision
The Rand value is based on a standardized subsidy amount, not on actual expenditure.

Vulnerability 3: Data Source Type
The risk of using standardized proxy data for the subsidies instead of actual amounts has not been established.

Recommendations for improvement:
Recommendation 1: Validity
The partner should be requested to report the Rand Value

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX I: DQA MUNICIPAL INFRASTRUCTURE INVESTMENT UNIT

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

MUNICIPAL INFRASTRUCTURE INVESTMENT UNIT (MIIU)
Check-sheet 1 of 1

Strategic Objective	**:**	No. 6 - Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	IR 6.2 - Previously ineligible households, developers, builders and municipal service providers obtaining access to credit
Performance Indicator	**:**	PI 6.2.1 - Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision (leveraged funds by definition).
Contractual Obligations	**:**	"… these measure the Rands leveraged and the households served under programs supported by the MIIU. These can be examined through USAID/SA's website …" (page 12 of the contract)
Data Source(s)	**:**	Letters/e-mails from independent consultants for the estimated total value of the projects.
Year or Period for which the data are being reported	**:**	1998 to 2002
Date(s) of Assessment	**:**	21 – 22 August 2003
Location(s) of Assessment	**:**	MIIU offices – Midrand Dr. James Leigland, Municipal Infrastructure Specialist Jackie Lesaoane, Project Administrator James Dohrman, Municipal Infrastructure Specialist
Assessment team members	**:**	On-site was Jacqui van Graan and Richard Martin Consolidator – Dr Penelope Richards

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:
_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.85]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	There is a direct relationship between the contracts MIIU are linkage officers for and the leveraged Rand value that is to be reported. The monies involved in these contracts are for the total build of the houses including the infrastructure.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	NOT APPLICABLE – no survey involved.
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	Data is obtained from letters/e-mails from independent consultants who calculate the estimated value of the projects. These Rand values are reported to USAID/SA as estimated nominal Rand values.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	Consultants use an array of formulae to calculate the estimated values – not incentive based.
➢ Are definitions for data to be collected operationally precise?	❑	✓	**1**	Because actual values cannot be reported estimated nominal values are reported, bringing error into the data reported. No calculations were seen therefore cannot determine the operational preciseness of the definitions. See Non-Conformity 1 below.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	NOT APPLICABLE to MIIU – data obtained from other sources.
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	
Transcription Error				
➢ What is the data transcription process?	✓	❑	**3**	Jackie Lesaoane captures the data into the spreadsheet and Dr. James Leigland or James Dohrman checks the data capture before reporting to USAID/SA.
➢ Is there potential for error?	❑	✓	**3**	LOW transcription error potential – data capture checked before reporting.
➢ Are steps being taken to limit transcription error?	✓	❑	**3**	Data capture is checked before reporting.
➢ Have data errors been tracked to	❑	❑	**N/A**	No back tracking of data errors were

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.85]				
	Yes	**No**	**Score**	**Comments**
their original source and mistakes corrected?				found at audit but errors are corrected when found.
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator?	✓	❑	**3**	Contracts generally refer to future expenditure and intention to undertaken specified and unspecified capital expenditure. This data must be manipulated to arrive at the gross value of the contract to be reported.
➢ Is there manipulation of secondary and/or tertiary data?	❑	✓	**3**	Data as obtained from the consultants and those directly from contracts are reported directly – Sometimes secondary and tertiary sources are used and must be manipulated. No complicated manipulation of data by MIIU themselves and by using experienced consultants to perform the calculations.
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	
➢ Are the correct formulae being applied?	❑	❑	**N/A**	It is arguable whether signature of a contract is leveraging funds within the year of reporting, or obtaining a promise to leverage the funds at a future date
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**N/A**	
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	Could not find evidence of missing data at audit.
➢ Are final numbers reported accurate? (e.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	All totals checked, added up correctly in so far as the arithmetical calculations are concerned.
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	✓	❑	**3**	No sampling – the sample equals the population.
➢ Did all units of the population have an equal chance of being selected for the sample?	✓	❑	**3**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	No sampling involved.
➢ Is the sample of adequate size?	✓	❑	**3**	No sampling – the sample equals the population.
➢ Are the data complete? (i.e., have all data points been recorded?)	✓	❑	**3**	No missing data were found at time of audit. No evidence at audit to suggest otherwise.

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.12]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	The data is always transcribed/captured from the data sources into the electronic spreadsheet.
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	Data is taken from contracts or letters/e-mails of consultants. The data source is the instrument.
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	❑	✓	**1**	Rand values on contracts are set by the contract. Estimated project Rand values are calculated via an intricate set of formulae that questions the reliability of the values . However, there are substantial variations in how the data are derived. It is impossible to standardize due to variations in the nature of contracts, for example management contracts and capital expenditure contracts. See Non-Conformity 2 below.
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	NOT APPLICABLE – no sampling involved.
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**3**	Data capture into the spreadsheet is checked.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	❑	✓	**1**	No periodic review of data collection is actioned. MIIU have no contact with the parties involved in the contracts after they have signed the contracts however when contracts are renegotiated, data are revised. See Non-Conformity 3 below.
➢ Do these procedures provide for periodic sampling and quality assessment of data?	❑	❑	**N/A**	NOT APPLICABLE – no procedures in place.
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**2**	No procedures in writing but unwritten process, e.g. Jackie Lesaoane is responsible for data capture and the follow up of costs directly to MIIU. See Non-Conformity 4 below.
➢ Are data problems at each level reported to the next level?	✓	❑	**3**	Jackie Lesaoane reports everything to Dr. James Leigland.
➢ Are data quality problems clearly described in final reports?	❑	✓	**1**	There is no narrative in the quarterly reports – purely a table of data that was designed by MIIU for purposes of reporting to USAID/SA. See Non-Conformity 5 below.

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.80]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Data are reported on a quarterly basis.
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	**N/A**	Incidental data collection.
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	Data is reported on a quarterly basis.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	The data table designed by MIIU for reporting purposes to USAID/SA is from the inception of MIIU, i.e. 1998. All projects from 1998 are reported with new projects added as they are signed. See Vulnerability 1 below.
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	Data is reported on a quarterly basis.
➢ Is the date of collection clearly identified in the report?	❑	✓	**2**	Only the years are reported, i.e. 1998 - 2002. See Non-Conformity 6 below.

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 1.00]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	The data is fundamentally predictive in that it reflects a contractual commitment to spend funds. Errors can occur due to unforeseeable circumstances.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**1**	Margin of error is not mentioned at all. The report is purely a table of data. See Non-Conformity 7 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**1**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	❑	**1**	UNKNOWN – because the validity and reliability of the tertiary data cannot be determined there is an absolute risk involved. See Non-Conformity 7 below.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	✓	❑	**3**	MIIU only facilitates contracts between parties so manipulation for "personal" gain would be irrelevant. Primary data is derived from written contracts. Aspects of the data, e.g. proportion of HDP, are assumed without validation. There is potential incentive to exaggerate the

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
				figures
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	MIIU are "linkage officers" and should remain objective when drawing up these contracts between parties. There is no audit evidence to prove otherwise. The objectivity of the consultants is unknown and could not be audited at time of audit. Yes, in that published data is used for population and household size, but attribution of the numbers served is very imprecise.
➢ Has there been independent review?	✓	❑	**3**	USAID/SA visited on 08 May 2002 – "reasonably valid data".

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 1.00]				
What is the main source of data? (Tick the appropriate box below)				
Primary		Secondary		Tertiary		Mixed	✓
	Yes	**No**	**Score**	**Comments**
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	✓	❑	**1**	Numbers receiving *new* services and numbers receiving *improved* services are aggregated. Data obtained from contracts and calculated estimates from consultants are aggregated but also reported separately for each project on the data table used for reporting purposes. See Non-Conformity 8 below.
➢ Have the risks associated with secondary and tertiary data sources been identified?	✓	❑	**1**	MIIU are aware of the risk involved in reporting the calculated estimated nominal values and has apparently spoken to USAID/SA in this regard. Due to the length of the contracts it is the only option open to them to be able to report something. See Non-Conformity 8 below.
➢ Has the credibility of the secondary and tertiary data been established?	❑	✓	**1**	The financial model to do the estimated calculations is very complicated and therefore it is out-sourced to independent consultants. These estimated Rand values introduce a multiple of errors besides the fact that it is an estimate in the first place. See Non-Conformity 8 below.

Non-Conformities:

Non-Conformity 1	Non Sampling Error: MAJOR – Although it has been determined that USAID/SA would prefer partners to report actual Rand values, MIIU report estimated nominal Rand values. The indicator definition does not specify this distinction. The operational preciseness of the definitions cannot be determined, introducing an absolute risk to data being reported.
Non-Conformity 2	Consistency: MAJOR – Although MIIU collect data as directly from the contracts or from letters/e-mails of independent consultants, the reliability of the data from these independent

	consultants is a risk – there are too many intricate calculations that introduce increasing margins of error with each calculation. This questions the reliability of these estimated Rand values.
Non-Conformity 3	Internal Quality Control: MAJOR – MIIU cannot report actual leveraged Rand values because they have no contact with the parties of the contract after signage. There is no way the reported values can be checked for accuracy.
Non-Conformity 4	Transparency: MINOR – There is no periodic review of data collection, maintenance and processing. MIIU have no further contact with the parties of the contracts after signing unless a contract is re-negotiated.
Non-Conformity 5	Transparency: MAJOR – Data quality problems are not mentioned at all in the reports seen. Due to the unreliability of the estimated values this is a significant risk.
Non-Conformity 6	Currency: MINOR – The period the data was collected cannot be determined exactly – giving only a year is not sufficient a time period. Important data can be excluded/included if the exact dates (at least months) are not reported.
Non-Conformity 7	Precision: MAJOR – No margin of error has been determined and because MIIU does not obtain actual Rand value data, this cannot be measured in the present process leaving the system open to absolute risk.
Non-Conformity 8	Data Source Type: MAJOR – The data provided by the tertiary data sources – independent consultants – use many intricate calculations that introduce increasing margins of error with each calculation. There is an inherent risk of reporting estimated values but values based on an intricate set of calculations is even more risky due to this error factor increase.
Strengths: Nil.	
Vulnerabilities:	
Vulnerability 1	Currency: All projects from inception in 1998 are reported over and over with each quarterly report. There is a possibility that these cumulative amounts are reported over and over to Washington thus reporting elevated incorrect results.
Recommendations for improvement (R):	
Recommendation 1	Reliability: Encourage the partners to report on any issues that may impact on data quality.
Recommendation 2	Currency: Encourage partners to include the exact time period of their report on their report.
Recommendation 3	Precision: MIIU should report actual Rand values unless USAID/SA are content and aware of the inherent risk involved in reporting estimated data.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX J: DQA KWA-ZULU NATAL PROJECT PREPARATION TRUST

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

PROJECT PREPARATION TRUST OF KWAZULU NATAL (PPTKN)
Check-sheet 1 of 1

Strategic Objective	**:**	MM - The promotion of a more integrated and sustainable approach to housing focusing specifically on pilot projects associated economic development and HIV/AIDS relief. Team - Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	MM - The preparation of between four and six housing projects associated either with HIV/AIDS relief or economic development, the acquisition of development funds and other support for the projects, and the commencement of project implementation. Team - -
Performance Indicator	**:**	MM - The approval of R42Million for 1608 households (primary quantitative indicator only). Note that the self-assessment / scoring below is focused on these primary indicators. Team: SO 6.1(e) - Rand value of municipal services – housing **AND** SO 6.2(e) - Number of households receiving services – housing
Contractual Obligations	**:**	SO 6.1(e) and PI 6.1.3
Data Source(s)	**:**	MM - Resolutions of approved funds from the Department of housing and other funders. Team - Subsidy approvals from DoH
Year or Period for which the data are being reported	**:**	MM - 01st February 2003 to 31st August 2004. Team – Stage 2 of Implementation Plan – Pre-feasibilities (report dated 19 August 2003). The completion of Stage 2 was 31 July 2003.
Date(s) of Assessment	**:**	MM - Assessment is ongoing and results are written up at the end of each reporting period to USAID (ie: 31/03/03; 31/07/03; 31/12/03; 31/07/04; 30/09/04) Team - 26 August 2003
Location(s) of Assessment	**:**	PPTKN offices – Liberty House, Durban Mark Misselhorn, CEO
Assessment team members	**:**	MM - PPT's CEO assisted by PPT's Finance and Office Administrator and project managers Mike Fraser and Vusi Ngwenya. Results are reported both the USAID via GMAC and to PPT's Board of Trustees. Team - On-site was Jacqui van Graan Consolidators – Richard Martin and Dr Penelope Richards

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

MM comment on DQA Checklist -
NOTE: In many cases questions have not been answered or scored because they do not appear to be relevant to the PPT assessment and data collection process, which is very simple and straightforward. Questions have only been answered where we are able to provide a relevant response, which is only in some of the cases. Feedback from Khulisa may assist PPT in providing further feedback.

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.85]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	❑	✓	**2**	MM – Yes since housing and infrastructure is the core of the project, the link is solid. However the associated economic development and HIV/AIDS relief is not as well correlated to the primary indicator. Whilst the primary indicator does provide a means of measurement, the criteria for assessment of success in these areas are somewhat more complex. The development of criteria to assess these projects will be undertaken as part of the program work PPT is undertaking. Team – The data PPTKN report are combined Rand values for the building of the house in total including the entire infrastructure. That which is reported does not relate directly to these two indicators because these two indicators are for housing only – this will result in inflated results. See Non-Conformity 1 below.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	MM - Not applicable – the data source is not a survey. Team – NOT APPLICABLE – no survey involved.
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	MM - The data is in the form of resolutions or other written evidence of funding approvals. No special instrument is therefore required. Team – No separate instrument – the data source is the instrument.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	Subsidy approvals are donation and income based.
➢ Are definitions for data to be collected operationally precise?	❑	✓	**2**	The Rand value is calculated from the number of houses that will be built – a set amount is subsidised for each house.

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.85]				
	Yes	**No**	**Score**	**Comments**
				See Non-Conformity 1 below.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	NOT APPLICABLE – no enumerators involved.
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	
Transcription Error				
➢ What is the data transcription process?	✓	❑	**3**	The Project Manager captures the data in a spreadsheet; MM or the finance and office administrator verifies the capture in the spreadsheet. MM types report to USAID/SA himself. MM verifies the secondary data from DoH – the approved subsidy – with the spreadsheet.
➢ Is there a potential for error?	❑	✓	**3**	LOW transcription error potential – double checks are made to avoid errors.
➢ Are steps being taken to limit transcription error?	✓	❑	**3**	Double checks are made to avoid errors.
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	No evidence of data errors were found at audit.
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator:	❑	❑	**N/A**	NOT APPLICABLE - Only secondary data involved.
➢ Is there manipulation of secondary and/or tertiary data?	❑	✓	**3**	MM – No. Team – The data reported is obtained directly from the approved subsidy document but MM does check the calculations himself. The Rand value is calculated from the number of houses to be built and the subsidy amount per house is the same value (simple multiplication).
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	NOT APPLICABLE - no data manipulation.
➢ Are the correct formulae being applied?	❑	❑	**N/A**	
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**N/A**	
➢ Have procedures for dealing with missing data been correctly applied?	✓	❑	**N/A**	Team – No evidence of missing data was found at audit.
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	Numbers reported in the 19 August 2003 report were verified as accurate.
Representativeness of Data				

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.85]				
	Yes	**No**	**Score**	**Comments**
➢ Is the sample from which the data are drawn representative of the population served by the activity?	✓	❑	**3**	No sampling involved – the sample equals the population.
➢ Did all units of the population have an equal chance of being selected for the sample?	✓	❑	**3**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	No sampling involved.
➢ Is the sample of adequate size?	✓	❑	**3**	No sampling involved – the sample equals the population.
➢ Are the data complete? (i.e., have all data points been recorded?)	✓	❑	**3**	All data points available at the time of report were reported – only the "approved in-principle" document had been received for 1 of the 10 projects. No evidence of missing data was found at audit.

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	MM – Yes - the same source of data is always used although there is a range of funders, so some variation occurs between different funders in terms of the way they approve funding. Team – Although this agreement has only started, this statement has been made based on the data reported for the first agreement direct with USAID/SA – the resolutions are used for data collection without exception.
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	Although this agreement has only started, this statement has been made based on the data reported for the first contract direct with USAID/SA – the resolutions are used for data collection without exception. The data source is the instrument.
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	❑	❑	**N/A**	MM – Yes. Team – NOT APPLICABLE – only a single secondary data source – DoH.
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	MM – Yes. Team – NOT APPLICABLE – no sampling involved.
Internal quality control				
➢ Are there procedures to ensure	✓	❑	**3**	MM –Yes - obtaining approvals in written

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
that data are free of significant error and that bias is not introduced?				form is usually sufficient to ensure that the approved figures are correct, this is especially so since the figures relate back to budgets and estimates in the feasibilities / business plans that PPT has prepared – this therefore provides an additional source of verification. Team – Crosschecks of all data capture are all part of their process. In fact the data from the source itself is also checked against the initial subsidy applications. Any extra funds are also followed up and entered into the spreadsheet but this is usually beyond the grant period.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	MM – Yes - PPT follows up with funders from time to time to see if additional funds have been approved. Team – Once the resolution of approval has been given and implementation starts, data is verified once the detailed design is complete and the town plan approved by the municipality. Data is again verified once the township is established and again as built. All these verifications of data usually occur after the grant has been closed out.
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	✓	❑	**3**	MM – not required. Team – PPTKN have a documented code of practice (did not physically see this document as MM had to search for it off their network). All data collected is reported – no cleaning of data relevant. Intent is proven.
➢ Are data problems at each level reported to the next level?	❑	❑	**N/A**	MM (CEO) is the highest level at PPTKN.
➢ Are data quality problems clearly described in final reports?	❑	❑	**N/A**	No evidence of data errors was found at audit. None to report at present – so could not be audited at the time of audit.

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	MM – Yes - although occasionally there can be delays in obtaining written confirmation from funders. Team – A narrative report will be handed in on a quarterly basis. Reports are also handed in against the milestones in the agreement.
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	**N/A**	MM – Yes - accurate schedules are kept. Team – NOT APPLICABLE – incidental data collection. However PPTKN have a process by which follow up data is

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
				checked on a weekly basis for inclusion into the spreadsheet.

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	MM – Yes - reporting takes place both against USAID timetable and PPT's own system of trimesterly reporting. Team – Reports are handed in to USAID/SA every quarter.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	MM – Yes. Team – The grant started on 23 February 2003 and the last milestone report is dated 19 August 2003.
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	MM – Yes - although there can be delays in collection due to funders being slow in releasing data. Team – Reporting on indicators is quarterly as well as with milestone completion reports.
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	MM – Yes. Team – The milestone report is dated 19 August 2003 and states that it is the "Report on Stage 2 of the Implementation Plan – Pre-feasibilities" – this implies when referring back to the agreement which time period is being referred to. End date of Stage 2 stated as 31 July 2003 in report. (tables within the report state the time period of interest by date !!)

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	Team – No margin of error calculated.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**2**	See Non-Conformity 2 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	✓	**3**	Crosschecks make for a sound system.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or	✓	❑	**3**	MM – Yes. Team – PPTKN only facilitate the implementation process and more than 1

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
personal reasons?				person completes verifications.
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	MM – Yes - data is verified by more than one person. Team - PPTKN only facilitate the implementation process and more than 1 person completes verifications.
➢ Has there been independent review?	✓	❑	**3**	USAID/SA have visited the site but no written confirmation has been received. There were verified visits from USAID/SA on the first contract with USAID/SA. See Vulnerability 1 below.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3.00]								
What is the main source of data? (Tick the appropriate box below)								
Primary		Secondary	✓	Tertiary		Mixed		
	Yes	**No**	**Score**	**Comments**				
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	MM – Yes. Team – Only secondary data involved.				
➢ Have the risks associated with secondary and tertiary data sources been identified?	✓	❑	**3**	MM verifies the approved subsidies against the initial subsidy application – see the procedure for periodic review under Internal Quality Control with Reliability.				
➢ Has the credibility of the secondary and tertiary data been established?	✓	❑	**3**	Resolutions are accepted as a credible data source.				

Non-Conformities:	
Non-Conformity 1	Face Validity: MINOR – PPTKN are reporting directly from the bulk subsidy approval they receive from DoH. The contract is to build x number of houses and the Rand value is determined by a simple multiplication as each subsidy for each house is exactly the same. The risk lies with the fact that the subsidy includes the infrastructure and is therefore not just for housing as is defined by the indicators.
Non-Conformity 2	Precision: MINOR – the margin of error has not been determined.
Strengths:	
Strength 1	The simplicity of the data management system makes verification easy.
Vulnerabilities:	
Vulnerability 1	Integrity: No written records on visits to PPTKN were seen at audit.
Recommendations for improvement (R):	
Recommendation 1	Ensure that written records are kept of all visits to all partners.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX L: DQA JOHANNESBURG HOUSING COMPANY

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

Partner: Johannesburg Housing Company
Check-sheet 1 of 2

Strategic Objective:	SO6: Increased access to shelter and environmentally sound municipal services
Intermediate Result:	IR 6.2.2:Number of households assisted to obtain shelter/urban services through the provision of credit and subsidies to low income communities
Performance indicator:	2 (e) Number of households receiving services – housing
Indicator description:	Number of HDP households assisted to obtain new or improved shelter/urban services through the provision of credit and subsidies
Contractual Obligations:	Annual report
Data source(s):	Report October 2002
Year or period for which the data are being reported:	July 2001 – June 2002
Date(s) of assessment:	27 August 2003
Location(s) of assessment:	Offices of the JHC, Johannesburg
Assessment team members:	Richard Martin

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score =2.60]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	Number of shelter units produced is the primary indicator of the program which is a component of the national housing delivery system
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	Not applicable
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**1**	Data collection is efficient, but numbers are inflated by inclusion of projects completed prior to USAID assistance. See non-conformity 1.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Data is collected via project management and accounting systems – no special enumerators involved
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	Not applicable
Transcription Error				
➢ What is the data transcription process? (Level of risk?)	❑	✓	**3**	Data is aggregation of accounting records by computer: effectively no transcription used
➢ Is there potential for error?	❑	✓	**3**	There is no likelihood of error for current data
➢ Are steps being taken to limit transcription error?	❑	❑	**N/A**	Not applicable
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	Not applicable
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator:	❑	✓	**3**	
➢ Is there manipulation of secondary and/or tertiary data?	❑	✓	**3**	
➢ How are the risks associated with manipulating data identified	❑	❑	**N/A**	Not applicable

1. VALIDITY—Do the data adequately represent performance? [Average Score =2.60]				
	Yes	**No**	**Score**	**Comments**
and managed?				
➢ Are the correct formulae being applied?	❑	❑	**N/A**	Not applicable
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	✓	❑	**3**	The same data collection protocols are applied.
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	Not applicable
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	❑	✓	**1**	See non-conformity 1
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑	**N/A**	Not applicable – no sampling
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑	**N/A**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	
➢ Is the sample of adequate size?	❑	❑	**N/A**	
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑	**N/A**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =3.00]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**3**	
➢ Is the same sampling method used from year to year, location to location, data source to data	❑	❑	**N/A**	Not applicable – no sampling

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =3.00]				
	Yes	**No**	**Score**	**Comments**
source?				

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =3.00]				
	Yes	**No**	**Score**	**Comments**
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	❑	❑	**3**	The nature of the data and the fact that it is 100% sample means that there is no risk of bias.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	
➢ Do these procedures provide for periodic sampling and quality assessment of data?	❑	❑	**N/A**	Not applicable
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**3**	The project documentation system is central to the records of the Company Procedures Manual.
➢ Are data problems at each level reported to the next level?	❑	✓	**N/A**	Not applicable
➢ Are data quality problems clearly described in final reports?	❑	✓	**N/A**	Not necessary as they resolved at the point of entry

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**2**	Data are available from a live database, but were only reported to USAID annually. Note: a quarterly report was produced but did not report on specific indicators. See non-conformity 2.
➢ Is a regularized schedule of data collection in place to meet program management needs?	✓	❑	**2**	
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	As noted above: data are available from a live data base. However, the annual reporting period used is different from the USAID fiscal year. A potential problem exists in that respect.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	❑	✓	**1**	Number of units reported on a cumulative basis, and includes units completed before USAID support was provided. See non-conformity 1.
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 3.00]

	Yes	No	Score	Comments
➢ Is the margin of error less than the expected change being measured?	✓	❑	**3**	Primary data used, so no margin of error permitted
➢ Is the margin of error acceptable given the likely management decisions to be affected?	✓	❑	**3**	
➢ Have targets been set for the acceptable margin of error?	❑	❑	**N/A**	Not applicable
➢ Has the margin of error been reported along with the data?	❑	❑	**N/A**	Not applicable
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	❑	**N/A**	Not applicable

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]

	Yes	No	Score	Comments
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	❑	❑	**N/A**	Not applicable
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	Basic nature of the data ensures objectivity
➢ Has there been independent review?	✓	❑	**3**	Audits and evaluation by other donors

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3.00]

What is the main source of data? (Tick the appropriate box below)

Primary	✓	Secondary		Tertiary		Mixed	

	Yes	No	Score	Comments
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	
➢ Have the risks associated with secondary and tertiary data sources been identified?	✓	❑	**3**	Primary data are used
➢ Has the credibility of the secondary and tertiary data been established?	❑	❑	**N/A**	Not applicable

Non-conformities:
Non-conformity 1: Validity
MAJOR - Data is aggregated from years prior to the grant agreement thus creating a major fault.

Non-conformity 2: Timeliness
MINOR - Data are available from a live database, but were only reported to USAID annually. Note: a quarterly report was produced but did not report on specific indicators. See non-conformity 2

Strengths and Vulnerabilities:
Strength 1: Validity
The unit of measurement is well defined and represents the focus of the program

Strength 2: Reliability
The data come from a well-audited data base

Strength 3: Precision
No sampling is involved

Strength 4: Integrity
Data is cross checked against auditable facts

Vulnerabilities:
Vulnerability 1: Timeliness
The annual reporting period used is different from the USAID fiscal year.

Recommendations for improvement:
Since the program is completed recommendations are not required

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

Partner: Johannesburg Housing Company
Check-sheet 2 of 2

Strategic Objective:	SO6: Increased access to shelter and environmentally sound municipal services
Intermediate Result:	IR 6.2.1 Rand value of credit and subsidies obtained from households for HDP shelter and service provision
Performance indicator:	1 (e) Rand Value of municipal services completed - housing
Indicator description:	Total Rands in millions provided, including funds for new or improved housing or services leveraged for HDP households
Contractual Obligations:	Annual report
Data source(s):	Report October 2002
Year or period for which the data are being reported:	July 2001 – June 2002
Date(s) of assessment:	27 August 2003
Location(s) of assessment:	Offices of the JHC, Johannesburg
Assessment team members:	Richard Martin

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score =3.00]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	The number of shelter units produced is the primary indicator of the program which is a component of the national housing delivery system: Rand value is derived directly from that
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	Not applicable
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	Accounts records used, no likelihood or error
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Data is collected via project management and accounting systems – no special enumerators involved
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	Not applicable
Transcription Error				
➢ What is the data transcription process? (low level of risk)	✓	❑	**3**	Data is aggregated manually from accounts, potential for error
➢ Low potential for error?	✓	❑	**3**	Zero
➢ Are steps being taken to limit transcription error?	❑	❑	**N/A**	Grant period has expired – no further action contemplated
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	Not applicable
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator:	❑	✓	**3**	The data are derived from accounts, and do not require manipulation
➢ Is there manipulation of secondary and/or tertiary data?	❑	✓	**3**	
➢ How are the risks associated with manipulating data identified and managed?	❑	✓	**3**	No specific risks are identified
➢ Are the correct formulae being applied?	❑	❑	**N/A**	No formulae are used

1. VALIDITY—Do the data adequately represent performance? [Average Score =3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	✓	❑	**3**	The same data collection protocols are applied.
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	Not applicable
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑	**N/A**	Not applicable – no sampling
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑	**N/A**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	
➢ Is the sample of adequate size?	❑	❑	**N/A**	
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑	**N/A**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =3.00]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	❑	✓	**3**	Initially an Attachment 8 form was used. The next two years data was submitted by e-mail to USAID Pretoria. No copies of these submissions are available
➢ Is the same instrument used to collect data from year to year, location to location?	❑	✓	**3**	See above
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	❑	❑	**N/A**	Not applicable
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	Not applicable – no sampling
Internal quality control				
➢ Are there procedures to ensure	❑	❑	**3**	The nature of the data and the fact that it

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =3.00]				
	Yes	**No**	**Score**	**Comments**
that data are free of significant error and that bias is not introduced?				is 100% sample means that there is no risk of bias.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	The same data are used as in financial reporting, e.g. Annual Report
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**3**	Data taken from accounts allocated in accordance with Company Procedures Manual
➢ Are data problems at each level reported to the next level?	❑	❑	**N/A**	Not applicable
➢ Are data quality problems clearly described in final reports?	❑	❑	**N/A**	Not applicable

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Data are available from accounts, which are maintained on a continuous basis.
➢ Is a regularized schedule of data collection in place to meet program management needs?	✓	❑	**3**	
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.25]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	✓	❑	**2**	Potential errors may occur in that a standard subsidy amount is used for reporting purposes. Actual levels of subsidy may differ. See non-conformity 1.
➢ Is the margin of error acceptable	✓	❑	**3**	The possible variations due to the above

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.25]				
	Yes	**No**	**Score**	**Comments**
given the likely management decisions to be affected?				error will not affect program management decisions

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.25]

	Yes	No	Score	Comments
➢ Have targets been set for the acceptable margin of error?	❑	❑		Not applicable
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	See non-conformity 1
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	✓	**2**	The amount could be obtained relatively cheaply. See non-conformity 1.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]

	Yes	No	Score	Comments
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	❑	❑	**N/A**	Not applicable
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	Objectivity of the primary data is ensured by financial audit procedures
➢ Has there been independent review?	✓	❑	**3**	Financial and performance audits undertaken by other donors

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 2.50]

What is the main source of data? (Tick the appropriate box below)

Primary		Secondary	✓	Tertiary		Mixed	

	Yes	No	Score	Comments
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	
➢ Have the risks associated with secondary and tertiary data sources been identified?	✓	❑	**2**	Attention has not been given to the potential error in using a standardized subsidy amount for reporting purposes. See non-conformity 1.
➢ Has the credibility of the secondary and tertiary data been established?	❑	❑	**N/A**	Not applicable

Non-conformities:
Non-conformity 1: Precision
MINOR - Potential errors may occur in that a standard subsidy amount is used for reporting purposes. Actual levels of subsidy may differ.

Strengths:
Strength 1: Validity
The data are based on well developed systems

Strength 2: Reliability
Data collection has followed a consistent and easily auditable process

Strength 3: Timeliness
Data are taken from a live data base

Vulnerabilities:
Vulnerability 1: Precision
The Rand value is based on a standardized subsidy amount, not on actual expenditure.

Vulnerability 2: Source of Data
The use of standardized subsidy amounts for reporting purposes could result in error.

Recommendations for improvement:
Since the program is completed recommendations are not required

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX K: DQA DEPARTMENT OF ENVIRONMENTAL AFFAIRS AND TOURISM / UNIVERSITY OF CAPE TOWN

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

ENIVRONMENTAL EVALUATION UNIT (EEU) – UNIVERSITY OF CAPE TOWN (UCT)
Check-sheet 1 of

Strategic Objective	**:**	Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	GCC Result 1 – Increased Participation in the UNFCCC
Performance Indicator	**:**	GCC 1.2 - Increased capacity to meet requirements of the UNFCCC, including activities in land use/forestry and energy/industrial/urban sectors
Contractual Obligations	**:**	No mention is made of indicators – only deliverables/milestones.
Data Source(s)	**:**	Training workshops/sessions attendance registers
Year or Period for which the data are being reported	**:**	June 1999 to March 2003
Date(s) of Assessment	**:**	27 August 2003
Location(s) of Assessment	**:**	EEU offices – UCT premises in Cape Town Dr Merle Sowman – Co-Director
Assessment team members	**:**	On-site was Jacqui van Graan Consolidators – Richard Martin and Dr Penelope Richards

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:
_____

Comments: _____

_____

_____

_____

_____

Key to Acronyms:
DrMS : Dr. Merle Sowman, Co-Director of EEU/UCT
Team : SO6 DQA Team

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.67]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	Provided that attending a workshop/ training activity can be equated to increased capacity. See Vulnerability 1 below.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	NOT APPLICABLE – no survey involved. The evaluation forms that were filled in by the candidates were extra research done by EEU for their own benefit.
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	✓	❑	**3**	The attendance registers have the basic details – name, organisation, address, telephone and facsimile numbers, e-mail address.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	Course/workshop attendees fill their own details in on the attendance register – a basic register to note the presence of each attendee.
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	A simple table of name and signature is self-explanatory.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Attendees fill in the attendance register themselves – no enumerators.
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	
Transcription Error				
➢ What is the data transcription process?	✓	❑	**3**	Person that manages the training session is responsible for the attendance register. Most times this person captures the register into the computer and the attendees check their personal details while on the training course – usually for the two-day courses there was not enough time for this. The course coordinator is responsible for capturing all data wrt courses into the database. No checks are made on this data capture.
➢ Is there a potential for error?	✓	❑	**2**	LOW transcription error potential - see Non-Conformity 1 below. In the example taken – Tshwane Municipality LA21 training course 26-27/11/2002 – one person on the written list is omitted from the report.
➢ Are steps being taken to limit	❑	✓	**2**	No checks are made on the data capture

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.67]				
	Yes	**No**	**Score**	**Comments**
transcription error?				of the course coordinator– see Non-Conformity 1 below.
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	✓	**2**	DrMS is not aware of any errors, as the reports are not distributed to the attendees. People that review the reports review the summative section and have no interest in the attendee detail. See Non-Conformity 1 and Vulnerability 2 below.
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator?	❑	✓	**3**	Only a simple total count of the attendees on the attendance register.
➢ Is there manipulation of secondary and/or tertiary data?	❑	❑	**N/A**	There is no secondary or tertiary data. Also no manipulation of data takes place.
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	
➢ Are the correct formulae being applied?	❑	❑	**N/A**	
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**N/A**	
➢ Have procedures for dealing with missing data been correctly applied?	❑	✓	**2**	EEU are not aware of missing data because the data capture by the course coordinator is not checked – see Non-Conformity 1.
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	All random total checks that were made were correct. Also the grant is complete and has been closed out.
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	✓	❑	**3**	The attendees that were counted represents the whole (100%) population. No sampling – the sample equals the population.
➢ Did all units of the population have an equal chance of being selected for the sample?	✓	❑	**3**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	No sampling involved.
➢ Is the sample of adequate size?	✓	❑	**3**	No sampling – the sample equals the population.
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	✓	**2**	There is missing data in the reports. But the grant has been closed out. See Non-Conformity 1 below.

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	Attendance registers filled in at training sessions are the only source of data.
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	Attendance registers only source of data.
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	❑	❑	**N/A**	NOT APPLICABLE – only one data source.
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	No sampling involved.
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	❑	✓	**2**	See Non-Conformity 1 below.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	❑	✓	**2**	See Non-Conformity 1 below.
➢ Do these procedures provide for periodic sampling and quality assessment of data?	❑	❑	**N/A**	No procedures in place.
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**2**	See Non-Conformity 1 below.
➢ Are data problems at each level reported to the next level?	❑	❑	**N/A**	Not auditable at time of audit – they are not aware of data problems. See Non-Conformity 1 below.
➢ Are data quality problems clearly described in final reports?	❑	✓	**2**	Reports make no reference to data problems at all. See Non-Conformity 2 below.

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Reports written after each training activity and kept until enough reports for minimum for tranche payments as per requirement of the contract.
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	**N/A**	Incidental data collection with each training session.
Currency				

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	Reports written after each training session and kept until enough reports for minimum for tranche payments as per requirement of the contract.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	
➢ Is the date of collection clearly identified in the report?	✓	❑	**3**	Exact dates and venue of the training session given on each report.

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 2.00]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	No margin of error determined because they do not have checks on data capture.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**2**	See Non-Conformity 3 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	✓	**2**	No crosschecks are made. See Non-Conformity 1 below.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	✓	❑	**3**	Attendees are either politicians or government officials and since reporting is done to USAID as well as DEAT, the information can be cross-checked by DEAT.
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	Attendees fill in the attendance register themselves.
➢ Has there been independent review?	✓	❑	**3**	USAID/SA (Sergio Guzman) apparently made regular visits but records of these were not kept. DrMS never received any formal visit DQA forms like other partners. See Vulnerability 4 below.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 3.00]							
What is the main source of data? (Tick the appropriate box below)							
Primary	✓	Secondary		Tertiary		Mixed	

	Yes	No	Score	Comments
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	Only primary data involved.
➢ Have the risks associated with secondary and tertiary data sources been identified?	❑	❑	**N/A**	Only primary data.
➢ Has the credibility of the secondary and tertiary data been established?	❑	❑	**N/A**	

Non-Conformities:	
Non-Conformity 1	Transcription Error: MINOR – There is no check on any capture of data and this has led to missing data as determined at audit. This has not been of concern until the time of the audit as they are more concerned with narrative – the evaluation part of the report.
Non-Conformity 2	Transparency: MINOR – Data quality problems are not mentioned at all in any of the reports.
Non-Conformity 3	Precision: MINOR – The margin of error has not been determined.
Strengths:	
Strength 1	Timeliness: A report is written and sent to USAID/SA after each training activity.
Vulnerabilities:	
Vulnerability 1	Face Validity: ONLY if attending a training activity can be equated to “increased capacity”, can there be a logical relationship between the activity and the indicator. However “increased capacity” refers to the actual increased understanding and way of thinking wrt to the training activity and this has not been measured.
Vulnerability 2	Transcription Error: The emphasis in the reports was placed on the narrative whereas USAID/SA places emphasis on numbers.
Vulnerability 3	Transcription Error: For reporting purposes, the number of attendees at training activities has been equated to “increased capacity”. See Vulnerability 1 for explanation.
Vulnerability 4	Integrity: Although DrMS stated that USAID/SA had visited DEAT/UCT a few times, no written record was found on-site at audit.
Recommendations for Improvement (R):	
Recommendation 1	Transparency: Encourage the partners to report on any issues that may impact on data quality.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX M: DQA MEGA-TECH

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

MEGA-TECH
Check-sheet 1 of 1

Strategic Objective	**:**	No. 6 - Increased access to shelter and environmentally sound municipal services.
Intermediate Result	**:**	-
Performance Indicator	**:**	Handling of Data from Partners.
Contractual Obligations	**:**	Handling of Data from Partners.
Data Source(s)	**:**	Partners
Year or Period for which the data are being reported	**:**	
Date(s) of Assessment	**:**	11 August 2003
Location(s) of Assessment	**:**	Mega-Tech offices – Brooklyn, Pretoria
Assessment team members	**:**	On-site was Richard Martin, Dr Penelope Richards and Jacqui van Graan

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:
_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.86]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	✓	❑	**3**	Primary function is to collect, collate and report data from contracted partners
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	NOT APPLICABLE – No survey involved.
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	❑	✓	**2**	Although the "USAID/GCC Grantee Data Collection Sheet" (GDCS) allows a standardized reporting system from all partners, it has room for improvement, e.g. the indicators number and the date of data collection does not appear on the sheet. See Non-Conformity 1 below.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	The "USAID/GCC Grantee Data Collection Sheet" has to be sent to MT along with a "GMAC Data Quality Questionnaire" – partners are made aware that their data will have to be verifiable.
➢ Are definitions for data to be collected operationally precise?	❑	✓	**2**	Although the partners were coached in how to fill in the GDCS, there are no explanations of definitions and there are problems with the GCC indicators as whole due to ambiguity. See Non-Conformity 2 below.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	✓	❑	**3**	The partners were coached in a half day workshop as to how to fill in the GDCS
➢ Were there efforts to reduce the potential for personal bias by enumerators?	✓	❑	**3**	The partners have to report their verifiable data – if the data cannot be verified they may lose their grant.
Transcription Error				
➢ What is the data transcription process?	❑	❑	**N/A**	MT handed in all the previous GDCS "as is" to USAID/SA – no transcription took place. See Vulnerability 1 below.
➢ Is there a potential for error?	❑	❑	**N/A**	
➢ Are steps being taken to limit transcription error?	❑	❑	**N/A**	
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.86]				
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator?	❑	❑	**N/A**	NOT APPLICABLE – MT only receives secondary and tertiary data. Also MT do not manipulate any data from partners – the GDCS were forwarded to USAID/SA "as is". See Vulnerability 1 below.
➢ Is there manipulation of secondary and/or tertiary data?	❑	❑	**N/A**	
➢ How are the risks associated with manipulating data identified and managed?	❑	❑	**N/A**	
➢ Are the correct formulae being applied?	❑	❑	**N/A**	
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	❑	❑	**N/A**	
➢ Have procedures for dealing with missing data been correctly applied?	❑	✓	**2**	Missing data is not always identified. See Non-Conformity 3 below.
➢ Are final numbers reported accurate? (e.g. does a number reported as a "total" actually add up?)	❑	✓	**2**	MT does not check all the GDCS before forwarding to USAID/SA, e.g. Cape Technikon (CT). See Non-Conformity 3 below.
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑	**N/A**	MT do not collect data as such, they merely "collate" data collected by the partners to forward to USAID/SA.
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑	**N/A**	
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	
➢ Is the sample of adequate size?	❑	❑	**N/A**	
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑	**N/A**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.75]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	❑	✓	**2**	Multiple changes from USAID/SA have resulted in multiple systems over the years. However MT have designed a standard form for all the partners that report to MT. See Non-Conformity 4 below.
➢ Is the same instrument used to collect data from year to year, location to location?	❑	✓	**2**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 2.75]				
	Yes	**No**	**Score**	**Comments**
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**3**	
➢ Is the same sampling method used from year to year, location to location, data source to data source?	❑	❑	**N/A**	MT does not collect data as such, they merely "collate" data collected by the partners to forward to USAID/SA.
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**3**	Informal systems that results in low auditability. MT does make on-site visits to partners that are contracted through them to USAID/SA but no written record is kept of these visits. See Vulnerability 2 and Vulnerability 3 below.
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	
➢ Do these procedures provide for periodic sampling and quality assessment of data?	✓	❑	**3**	
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	❑	✓	**3**	Informal procedures are in place and also MT merely collate collected data.
➢ Are data problems at each level reported to the next level?	✓	❑	**3**	Nomonde Mdhluli reports problems to Steve Horn.
➢ Are data quality problems clearly described in final reports?	❑	❑	**N/A**	MT does not collect data as such, they merely "collate" data collected by the partners to forward to USAID/SA – no reports are handed in to USAID/SA. See Vulnerability 1 below.

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.83]				
	Yes	**No**	**Score**	**Comments**
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	As far as MT involvement goes – MT has everything in place. The difficulty that is presented is that many of the smaller partners report prospective data rather than retrospective data. See Vulnerability 4 below.
➢ Is a regularized schedule of data collection in place to meet program management needs?	✓	❑	**3**	
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	As far as MT involvement goes – MT has everything in place. The difficulty that is presented is that many of the smaller partners report prospective data rather than retrospective data. See Vulnerability 4 below.
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 2.83]				
	Yes	**No**	**Score**	**Comments**
➢ Is the date of collection clearly identified in the report?	❑	✓	**2**	The GDCS given the partners by MT has no specified place for date(s) of collection. See Non-Conformity 1 below.

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 1.67]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	❑	**N/A**	No margin of error has been calculated by any of the partners that report through MT.
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	❑	**N/A**	
➢ Have targets been set for the acceptable margin of error?	❑	✓	**2**	No margin of error calculated therefore no target. See Non-Conformity 5 below.
➢ Has the margin of error been reported along with the data?	❑	✓	**2**	
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	❑	**1**	UNKNOWN – because the validity and reliability of the data reported to MT could not be verified at audit there is an absolute risk involved. MT is also responsible for data quality of the partners that report through them. See Non-Conformity 6 below.

5. INTEGRITY—Are data free of manipulation? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	✓	❑	**3**	Data is self-reported to MT via the GDCS. However there were no specific risks for manipulations of data identified at audit wrt MT – MT has nothing to gain by manipulating the data.
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	MT only collates the data collected by the partners – MT has nothing to gain by manipulating the data.
➢ Has there been independent review?	❑	✓	**2**	No written records of DQA's at MT were seen at audit. MT visits the partners contracted through them but they do not keep written records of these visits. See Non-Conformity 7 and Vulnerability 3 below.

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score =]								
What is the main source of data? (Tick the appropriate box below)								
Primary		Secondary		Tertiary		Mixed	✓	
	Yes	**No**	**Score**	**Comments**				

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score =]				
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	❑	**1**	UNKNOWN – partners reporting on the same indicators are aggregated and because the partners themselves were not audited it is unknown what data type they are reporting. See Non-Conformity 8 below.
➢ Have the risks associated with secondary and tertiary data sources been identified?	❑	✓	**1**	The data has not been verified because the partners themselves were not verified at audit. Also although on-site visits were made no DQA's were seen at audit. See Non-Conformities 1 through 8 below.
➢ Has the credibility of the secondary and tertiary data been established?	❑	✓	**1**	

Non-Conformities:	
Non-Conformity 1	Non Sampling Error: MINOR – The data collection instrument - "USAID/GCC Grantee Data Collection Sheet" – is somewhat complicated and is open to interpretation by reporting partners.
Non-Conformity 2	Non Sampling Error: MINOR – Definitions are not operationally precise.
Non-Conformity 3	Data Manipulation: MINOR – MT does not check all data contained in the GDCS before forwarding to USAID/SA, missing data and also the accuracy of totals is a risk.
Non-Conformity 4	Consistency: MINOR – The data collection system and/or instrument has been altered significantly during the reporting period. This introduces risk for collation of data at USAID/SA level.
Non-Conformity 5	Precision: MINOR – The margin of error has not been determined by any of the partners that report through MT.
Non-Conformity 6	Precision: MAJOR – It was not possible to determine the extent of error during this audit and thus it is not possible to determine whether such error can be reduced as a cost--effective and manageable interest.
Non-Conformity 7	Integrity: MINOR – MT have not had an on-site visit from USAID/SA – no written record was seen at audit.
Non-Conformity 8	Data Source Type: MAJOR– It is unknown if different data types are aggregated – this introduces an absolute risk.
Strengths:	
Strength 1	MT have tried to introduce a standardized data collection system for all the partners that report through them to USAID/SA – an idea that may be useful for all partners.
Vulnerabilities:	
Vulnerability 1	Transcription Error: Although no transcription took place during the previous run of data collection from MT, there is a possible risk that USAID/SA should be made aware of. USAID/SA has requested that MT collate the data reported – there may be a potential for error with the next data collection when MT collates all the data from the various partners due to inherent data differences.
Vulnerability 2	Internal Quality Control: The informal internal quality control systems at MT are not formalized resulting in low

	auditability
Vulnerability 3	Internal Quality Control: No written records are kept of on-site visits to partners.
Vulnerability 4	Currency: Potential error is introduced by partners which report a mix of prospective and retrospective data.
Recommendations for improvement (R):	
Recommendation 1	Non Sampling Error: Revise the GDCS to be fully representative of the indicators with the data collection period.
Recommendation 2	Non Sampling Error: Revise the definitions of the indicators so that these can be used on the GDCS as well and to remove ambiguities.
Recommendation 3	Consistency: Design and maintain a consistent data collection process / instrument.
Recommendation 4	Internal Quality Control: Formalize a good system that includes visits to partners, with written records, to verify the data reported to MT.
Recommendation 5	Precision: Encourage partners to have an auditable data trail for all data they report on.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

+

Assessment Team:

Mr. R. Martin (Team leader)

Dr. P.A. Richards

Mrs. J. van Graan

USAID SOUTH AFRICA SO6 DATA QUALITY ASSESSMENT

APPENDIX N: DQA FIRST RAND BANK

Submitted to:

USAID/South Africa

by

Khulisa Management Services (Pty) Ltd

and

Megatech

Solicitation Number 0112-0603-SOL-ME8

15 November 2003

DATA QUALITY ASSESSMENT CHECKLIST

Partner: FirstRand Bank
Check-sheet 1 of 2

Strategic Objective:	SO6: Increased access to shelter and environmentally sound municipal services
Intermediate Result:	IR 6.2.2: Number of households assisted to obtain shelter/urban services through the provision of credit and subsidies to low income communities
Indicator description:	Number of HDP households assisted to obtain new or improved shelter/urban services through the provision of credit and subsidies
Performance indicator:	6.2 (a) – (e) Number of households receiving municipal services
Contractual Obligations:	Annual report in prescribed format
Data source(s):	USAID Program Delivery Report, October 2002
Year or period for which the data are being reported:	April 31 2002 – March 31 2003
Date(s) of assessment:	28 August 2003
Location(s) of assessment:	Desk study of USAID Program Delivery Report, October 2002
Assessment team members:	Richard Martin

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.00]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	❑	✓	**1**	The data reported are an indirect product of the funds being spent. See non-conformity 1 and 2.
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	Not applicable
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	❑	✓	**1**	There is no specific instrument. Reliance is placed on loan agreements signed with borrowers. See non-conformity 1.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	Definition is accurate insofar as it refers only to funds spent. However, there are many potential measuring errors discussed below.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Not applicable
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	Not applicable
Transcription Error				
➢ What is the data transcription process? (low risk?)	✓	❑	**3**	Copied from loan agreements, and local government sources
➢ What is the potential for error? (low risk?)	✓	❑	**2**	Errors in using local government data are possible. See non-conformity 3.
➢ Are steps being taken to limit transcription error?	❑	✓	**2**	Not a major problem in light of the small volume of data
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	Not applicable
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator:	✓	❑	**1**	Yes there are several assumptions made in the manipulation of the data. See on-conformity 4
➢ Is there manipulation of secondary and/or tertiary data?	✓	❑	**1**	Sometimes secondary and tertiary sources are used and must be manipulated, e.g. local government municipal budgets and population data. See non-conformity 4.

1. VALIDITY—Do the data adequately represent performance? [Average Score = 2.00]

	Yes	No	Score	Comments
➢ Are the risks associated with manipulating data identified and managed?	✓	❑	**3**	The formulae to be used were agreed in advance with USAID
➢ Are the correct formulae being applied?	❑	✓	**1**	The values used in the formulae are such that there is no direct relationship between the funds spent and the results claimed. Attribution of the loan to specific households is extremely speculative, just as the division of the expenditure between different services. Non-conformity 5.
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	✓	❑	**3**	
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	Not applicable
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	In so far as the arithmetical calculations are concerned
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑	**N/A**	Not applicable
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑	**N/A**	Not applicable
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	Not applicable
➢ Is the sample of adequate size?	❑	❑	**N/A**	Not applicable
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑	**N/A**	Not applicable

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 3.00]

	Yes	No	Score	Comments
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score = 3.00]

	Yes	No	Score	Comments
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**3**	
➢ Is the same sampling method used from year to year, location to location, data source to data source?	✓	❑	**3**	
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**3**	The small amount of data makes this unnecessary
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	Not applicable: data are only reported at time of loan signature
➢ Do these procedures provide for periodic sampling and quality assessment of data?	❑	❑	**N/A**	Not applicable
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	✓	❑	**3**	Yes, as agreed with USAID
➢ Are data problems at each level reported to the next level?	❑	❑	**N/A**	Not applicable
➢ Are data quality problems clearly described in final reports?	❑	❑	**N/A**	Not applicable

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]

	Yes	No	Score	Comments
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Data are available on signature of a loan, but small number of loans means that program decisions are not made on the basis of the data
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	**N/A**	Not applicable
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	In view of the fact that the program consists of a few very large loans
➢ Is the date of collection clearly	✓	❑	**3**	

3. TIMELINESS—Are data collected frequently and are they current? [Average score = 3.00]				
	Yes	**No**	**Score**	**Comments**
identified in the report?				

4. PRECISION—Do the data have an acceptable margin of error? [Average score = 1.00]				
	Yes	**No**	**Score**	**Comments**
➢ Is the margin of error less than the expected change being measured?	❑	✓	**1**	It is not possible to establish a margin of error. The attribution of the expenditure to specific services and households cannot be verified. See non-conformity 6
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	✓	**1**	It is impossible to determine the margin of error as the data are not sufficiently verifiable
➢ Have targets been set for the acceptable margin of error?	❑	❑	**N/A**	Not applicable
➢ Has the margin of error been reported along with the data?	❑	❑	**N/A**	Not applicable
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	✓	❑	**1**	Only highly expensive and cumbersome data measurement could be used as a substitute. See non-conformity 6.

5. INTEGRITY—Are data free of manipulation? [Average score = 2.33]				
	Yes	**No**	**Score**	**Comments**
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	✓	❑	3	The formulae for the manipulation of data have been agreed with USAID, and calculations demonstrating adherence to the formulae are submitted with the primary data
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	3	Yes, in that verifiable expenditure is used as the primary data
➢ Has there been independent review?	❑	✓	1	

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 2.33]							
What is the main source of data? (Tick the appropriate box below)							
Primary		Secondary	✓	Tertiary		Mixed	✓

	Yes	**No**	**Score**	**Comments**
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	
➢ Have the risks associated with secondary and tertiary data sources been identified?	❑	✓	**2**	Formulae are based on socio-economic and population norms which are not verified in practice
➢ Has the credibility of the secondary and tertiary data been	✓	❑	**2**	Local government budgets and population data, therefore subject to rigorous

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score = 2.33]				
established?				checking

Non-conformities:
Non-conformity 1: Validity
Major - The formula used has two major weaknesses – while funds are being provided for general municipal use they are attributed to specific households and specific services. There is, in practice, no demonstrable link between the two.

Non-conformity 2: Validity
MINOR - There is no measurement of the degree to which the services provided are environmentally sustainable.

Non-conformity 3: Validity
MINOR - Errors in using local government data are possible.

Non-conformity 4: Validity
MAJOR - Significant risk is introduced with manipulations of secondary and tertiary data sources.

Non-conformity 5: Validity
MAJOR - The data are based on formulae and cost assumptions regarding the expenditure for specific services. Such cost data has not been verified in the field. It may be unduly conservative, but the fact has not been verified

Non-conformity 6: Precision
MAJOR - There are insufficient direct linked between the data and the results claimed to establish the margin of error

Strengths
Strength 1: Reliability
The data are consistently collected and applied in a transparent manner

Strengths 2: Timeliness
The data are available on an immediate basis.

Vulnerabilities:
Vulnerability 1: Data Source Type
The derivation of the population/household data has not been checked.

Recommendations for improvement (R):
Recommendation 1: Validity
The data should not be disaggregated into specific services unless the funds were specifically allocated for such a service.

GENERAL MANAGEMENT ASSISTANCE CONTRACT (GMAC)
Contract No: 674-C-00-01-10051-00

DATA QUALITY ASSESSMENT CHECKLIST

Partner: FirstRand Bank
Check-sheet 2 of 2

Strategic Objective:	SO6: Increased access to shelter and environmentally sound municipal services
Intermediate Result:	IR 6.2.1 Rand value of credit and subsidies obtained for households for HDP shelter and urban services provision
Performance indicator:	1 (a) – (d)
Indicator description:	Total Rands in millions provided, including funds for new or improved housing or services leveraged for HDP households
Contractual Obligations:	Annual report in prescribed form
Data source(s):	USAID Program Delivery Report, October 2002
Year or period for which the data are being reported:	April 31 2002 – March 31 2003
Date(s) of assessment:	28 August 2003
Location(s) of assessment:	Desk study of USAID Program Delivery Report, October 2002
Assessment team members:	Richard Martin

For Office Use Only

SO team leader approval:

X_____Date_____

Mission director or delegate approval:

X_____Date_____

Copies to:

_____

Comments: _____

_____

_____

_____

_____

1. VALIDITY—Do the data adequately represent performance? [Average Score =1.92]				
	Yes	**No**	**Score**	**Comments**
Face Validity				
➢ Is there a solid, logical relation between the activity or program and what is being measured?	❑	✓	**1**	The data reported are an indirect product of the funds being spent
Measurement Error				
***Sampling Error* (only applies when the data source is a survey)**				
➢ Were samples representative?	❑	❑	**N/A**	Not applicable
➢ Were the questions in the survey/questionnaire clear, direct, easy to understand?	❑	❑	**N/A**	
➢ If the instrument was self-reporting were adequate instructions provided?	❑	❑	**N/A**	
➢ Were response rates sufficiently large?	❑	❑	**N/A**	
➢ Has non-response rate been followed up?	❑	❑	**N/A**	
Non Sampling Error				
➢ Is the data collection instrument well designed?	❑	✓	**2**	There is no specific instrument. Reliance is placed on loan agreements signed with borrowers. See non-conformance 1.
➢ Were there incentives for respondents to give incomplete or untruthful information?	❑	✓	**3**	
➢ Are definitions for data to be collected operationally precise?	✓	❑	**3**	Definition is accurate insofar as it refers only to funds spent. However, there are many potential measuring errors discussed below.
➢ Are enumerators well trained? (How were they trained? Were they insiders or outsiders? Was there any quality control in the selection process?)	❑	❑	**N/A**	Not applicable
➢ Were there efforts to reduce the potential for personal bias by enumerators?	❑	❑	**N/A**	Not applicable
Transcription Error				
➢ What is the data transcription process?	❑	❑	**3**	Copied from loan agreements, and local government sources
➢ What is the potential for error?	❑	❑	**2**	Errors in using local government data are possible. See non-conformity 2.
➢ Are steps being taken to limit transcription error?	❑	✓	**2**	
➢ Have data errors been tracked to their original source and mistakes corrected?	❑	❑	**N/A**	Not applicable
Data Manipulation				
➢ Do primary data need to be manipulated to produce the data required for the indicator:	✓	❑	**1**	Yes there are several assumptions made in the manipulation of the data. There is no verifiable data to demonstrate that the expenditure reported made and identifiable change in the service provided. See non-conformity 3
➢ Is there manipulation of secondary and/or tertiary data?	✓	❑	**1**	Sometimes secondary and tertiary sources are used and must be manipulated, e.g. local government

1. VALIDITY—Do the data adequately represent performance? [Average Score =1.92]				
	Yes	**No**	**Score**	**Comments**
				municipal budgets and population data
➢ How are the risks associated with manipulating data identified and managed? (Risk?)	✓	❑	**2**	The formulae to be used were agreed in advance with USAID
➢ Are the correct formulae being applied?	❑	✓	**1**	The methodology cannot produce anything other than suggestive results with no direct attributability.
➢ Are the same formulae applied consistently from year to year, site-to-site, data source to data source?	✓	❑	**3**	
➢ Have procedures for dealing with missing data been correctly applied?	❑	❑	**N/A**	Not applicable
➢ Are final numbers reported accurate? (E.g., does a number reported as a "total" actually add up?)	✓	❑	**3**	In so far as the arithmetical calculations are concerned
Representativeness of Data				
➢ Is the sample from which the data are drawn representative of the population served by the activity?	❑	❑	**N/A**	Not applicable
➢ Did all units of the population have an equal chance of being selected for the sample?	❑	❑	**N/A**	Not applicable
➢ Is the sampling frame (i.e., the list of units in the target population) up to date? Comprehensive? Mutually exclusive (for geographic frames)	❑	❑	**N/A**	Not applicable
➢ Is the sample of adequate size?	❑	❑	**N/A**	Not applicable
➢ Are the data complete? (i.e., have all data points been recorded?)	❑	❑	**N/A**	Not applicable

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =2.85]				
	Yes	**No**	**Score**	**Comments**
Consistency				
➢ Is a consistent data collection process used from year to year, location-to-location, data source to data source (if data come from different sources)?	✓	❑	**3**	
➢ Is the same instrument used to collect data from year to year, location to location?	✓	❑	**3**	

2. RELIABILITY—Are data collection processes stable and consistent over time? [Average score =2.85]

	Yes	No	Score	Comments
➢ If data come from different sources are the instruments similar enough that the reliability of the data are not compromised?	✓	❑	**3**	
➢ Is the same sampling method used from year to year, location to location, data source to data source?	✓	❑	**3**	
Internal quality control				
➢ Are there procedures to ensure that data are free of significant error and that bias is not introduced?	✓	❑	**2**	The small amount of primary data makes this unnecessary. However, the data reported cannot be objectively verified
➢ Are there procedures in place for periodic review of data collection, maintenance, and processing?	✓	❑	**3**	Not applicable: data are only reported at time of loan signature
➢ Do these procedures provide for periodic sampling and quality assessment of data?	❑	❑	**N/A**	Not applicable
Transparency				
➢ Are data collection, cleaning, analysis, reporting, and quality assessment procedures documented in writing?	✓	❑	**3**	Yes, as agreed with USAID
➢ Are data problems at each level reported to the next level?	❑	❑	**N/A**	Not applicable
➢ Are data quality problems clearly described in final reports?	❑	❑	**N/A**	Not applicable

3. TIMELINESS—Are data collected frequently and are they current? [Average score =3.00]

	Yes	No	Score	Comments
Frequency				
➢ Are data available on a frequent enough basis to inform program management decisions?	✓	❑	**3**	Data are available on signature of a loan, but small number of loans means that program decisions are not made on the basis of the data
➢ Is a regularized schedule of data collection in place to meet program management needs?	❑	❑	**N/A**	Not applicable
Currency				
➢ Are the data reported in a given timeframe the most current practically available?	✓	❑	**3**	
➢ Are data from within the policy period of interest? (i.e., are data from a point in time after intervention has begun?)	✓	❑	**3**	
➢ Are the data reported as soon as possible after collection?	✓	❑	**3**	In view of the fact that the program consists of a few very large loans
➢ Is the date of collection clearly	✓	❑	**3**	

3. TIMELINESS—Are data collected frequently and are they current? [Average score =3.00]				
	Yes	**No**	**Score**	**Comments**
identified in the report?				

4. PRECISION—Do the data have an acceptable margin of error? [Average score =1.00]

	Yes	No	Score	Comments
➢ Is the margin of error less than the expected change being measured?	❑	❑	**1**	The only room for error lies in transcription and/or manipulation of local government budgetary and population information, however, interpretation is subject to major potential errors. The nature of the interpretation makes any verification impossible
➢ Is the margin of error acceptable given the likely management decisions to be affected?	❑	✓	**1**	The margin of error cannot be verified due to the nature of the data and the results.
➢ Have targets been set for the acceptable margin of error?	❑	❑	**N/A**	Not applicable
➢ Has the margin of error been reported along with the data?	❑	❑	**N/A**	Not applicable
➢ Would an increase in the degree of accuracy be more costly than the increased value of the information?	❑	❑	**N/A**	Not applicable

5. INTEGRITY—Are data free of manipulation? [Average score =2.33]

	Yes	No	Score	Comments
➢ Are mechanisms in place to reduce the possibility that data are manipulated for political or personal reasons?	❑	✓	**3**	The formulae for the manipulation of data have been agreed with USAID, and calculations demonstrating adherence to the formulae are submitted with the primary data
➢ Is there objectivity and independence in key data collection, management, and assessment procedures?	✓	❑	**3**	Yes, in that verifiable expenditure is used as a primary source
➢ Has there been independent review?	❑	✓	**1**	Nil demonstrated at audit

6. DATA SOURCE TYPE—What is the effect of the data type? [Average score =2.33]

What is the main source of data? (Tick the appropriate box below)

Primary		Secondary	✓	Tertiary		Mixed	✓

	Yes	No	Score	Comments
➢ Is there any aggregation of different data types into a single figure for reporting purposes?	❑	✓	**3**	
➢ Have the risks associated with secondary and tertiary data sources been identified?	❑	✓	**2**	Formulae depend on population and socio-economic norms which are not verified
➢ Has the credibility of the secondary and tertiary data been established?	✓	❑	**2**	Local government budgets and population data, therefore subject to rigorous checking

Non-conformities:
Non-conformity 1: Validity
MAJOR - There is no demonstrable link between the funds spent and the results claimed in terms of HDP households assisted to obtain shelter/urban services.

Non-conformity 2: Validity
MINOR - Errors in using local government data are possible.

Non-conformity 3: Validity
MAJOR - Significant risk is introduced with manipulations of secondary and tertiary data sources.

Non-conformity 4: Reliability
MINOR - The data reported cannot be objectively verified

Non-conformity 5: Precision
MAJOR - The margin of error cannot be demonstrated to be less than the expected change being measured as the nature of the interpretation makes any verification impossible

Strengths:
Strength 1: Timeliness
The data are available on an immediate basis

Vulnerabilities:
Vulnerability 1: Precision
There is no direct link between the expenditure claimed and the percentage of expenditure going to HDP households.